

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 84.8957 Seconds  
(without alignments)  
108.037 Million cell updates/sec

Title: US-09-932-322-1  
Perfect score: 44  
Sequence: 1 XXXCXPTGCGXXX 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : UniProt\_05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	81.8	192	2	Q98A9 RHILIO
2	36	81.8	360	2	Q86AK7 D1CDI
3	36	81.8	546	2	Q8KJ72 RHILIO
4	36	81.8	597	2	Q5AK44 D1CDI
5	36	81.8	601	2	Q7M4J3 D1CDI
6	36	81.8	608	2	Q627A0 CABBR
7	36	81.8	1710	2	Q54YG2 D1CDI
8	35	79.5	148	2	Q95KD2 MACFA
9	35	79.5	148	2	Q9N018 MACFA
10	35	79.5	370	2	Q5SBT6 D1CDI
11	35	79.5	370	2	Q5SBX0 D1CDI
12	35	79.5	466	2	Q5KKP3 CRYNE
13	35	79.5	571	2	Q9AQP5 PSERE
14	35	79.5	997	2	Q4SBN7 TETNG
15	35	79.5	1018	2	Q4SP52 TETNG
16	35	79.5	1839	2	Q4S795 TETNG
17	34	77.3	113	2	Q726J7 HUMAN
18	34	77.3	145	1	PA2X NOTSC
19	34	77.3	167	1	Q7U7V9 SYNPC
20	34	77.3	296	2	Q54YF0 D1CDI
21	34	77.3	312	2	Q9GTV3 ARATH
22	34	77.3	380	2	Q5AN14 D1CDI
23	34	77.3	465	2	Q6TVH9 GPDXV
24	34	77.3	487	2	Q4MNC1 ASPFU
25	34	77.3	504	2	Q5ZAZ8 ORYSA
26	34	77.3	507	2	Q5AN64 D1CDI
27	34	77.3	670	2	Q4SPD9 TETNG
28	34	77.3	870	2	Q5B3W6 EMBNI
29	34	77.3	932	2	Q54N02 D1CDI
30	34	77.3	998	2	Q869K4 D1CDI
31	34	77.3	999	2	Q55C92 D1CDI

32	34	77.3	1147	2	Q57XW5 GREYP	Q57XW5 trypanosoma
33	34	77.3	2217	2	Q6Y625 YVIRU	Q6Y625 iassa virus
34	33	75.0	77	2	Q4IR98 GIBRE	Q4IR98 gibberella
35	33	75.0	269	2	Q67W06 ORYSA	Q67W06 oryza sativ
36	33	75.0	330	2	Q5NIV4 9ARCD	Q5NIV4 uncultured
37	33	75.0	464	2	Q7XV21 ORYSA	Q7XV21 oryza sativ
38	33	75.0	1058	2	Q9LEB25 ARATH	Q9LEB25 arabidopsis
39	33	75.0	1468	2	Q80TF6 MOUSE	Q80TF6 mus musculus
40	32	72.7	64	2	Q92U25 RHIME	Q92U25 rhizobium m
41	32	72.7	123	2	Q926M4 LIRIN	Q926M4 listeria in
42	32	72.7	222	2	Q5NN00 ANOGA	Q5NN00 anopheles g
43	32	72.7	231	2	Q5TMZ9 ANOGA	Q5TMZ9 anopheles g
44	32	72.7	520	2	Q628K1 ORYSA	Q628K1 oryza sativ
45	32	72.7	1152	2	Q9FI26 ARATH	Q9FI26 arabidopsis
46	31	70.5	162	2	Q9WU17 MESAU	Q9WU17 mesocricetu
47	31	70.5	166	2	Q8T415 DROME	Q8T415 drosophila
48	31	70.5	183	2	Q4H6G8 GPBIO	Q4H6G8 deinococcus
49	31	70.5	229	2	Q4S7H4 TETNG	Q4S7H4 tetraodon n
50	31	70.5	239	2	Q93KV9 STRVR	Q93KV9 streptomyce
51	31	70.5	352	2	Q6VHM4 STRAM	Q6VHM4 streptomyce
52	31	70.5	361	2	Q8ECT9 SHEON	Q8ECT9 shewanella
53	31	70.5	482	2	Q9LZB8 ARATH	Q9LZB8 arabidopsis
54	31	70.5	615	2	Q22886 CAEEL	Q22886 caenorhabdi
55	31	70.5	719	2	Q54I92 D1CDI	Q54I92 dictyosteli
56	30	68.2	66	2	Q7N375 PHOLD	Q7N375 photornabdu
57	30	68.2	78	2	P90569 PLABE	P90569 plasmodium
58	30	68.2	95	2	P77130 ECOLI	P77130 escherichia
59	30	68.2	100	1	CHA3 BOMMO	P08929 bombyx mori
60	30	68.2	114	2	Q8WQ78 CAEEL	Q8WQ78 caenorhabdi
61	30	68.2	119	2	Q4UUP1 XANCP	Q4UUP1 xanthomonas
62	30	68.2	119	2	Q8P958 XANCP	Q8P958 xanthomonas
63	30	68.2	129	1	CHAI BOMMO	P08826 bombyx mori
64	30	68.2	132	1	CHAZ BOMMO	P08829 bombyx mori
65	30	68.2	133	2	Q9LIW4 ORYSA	Q9LIW4 oryza sativ
66	30	68.2	150	2	Q6D216 ERMCT	Q6D216 erwinia car
67	30	68.2	151	2	Q8ZD76 YERPS	Q8ZD76 yerstinia ps
68	30	68.2	151	2	Q66707 YERPS	Q66707 yerstinia ps
69	30	68.2	153	2	Q8KQC9 YERIN	Q8KQC9 yerstinia en
70	30	68.2	157	2	Q614N5 CABBR	Q614N5 caenorhabdi
71	30	68.2	171	2	Q9ULP0 RAT	Q9ULP0 rattus norv
72	30	68.2	175	2	Q6IJNO DROME	Q6IJNO drosophila
73	30	68.2	197	2	Q4IG50 9BROK	Q4IG50 burkholderi
74	30	68.2	197	2	Q4U1V1 9SMBG	Q4U1V1 punctatus p
75	30	68.2	199	2	Q7RUV7 NEUCR	Q7RUV7 neurospora
76	30	68.2	200	1	WIN1 SOLTU	P09167 solanum tub
77	30	68.2	201	2	003954 LYCBS	003954 lycopersico
78	30	68.2	203	2	Q75QH1 CAPCH	Q75QH1 capsicum ch
79	30	68.2	208	2	Q41230 TOBAC	Q41230 nicotiana t
80	30	68.2	210	2	Q9SEW3 CAPAN	Q9SEW3 capsicum an
81	30	68.2	211	1	WIN2 SOLTU	P09162 solanum tub
82	30	68.2	211	2	Q41231 TOBAC	Q41231 nicotiana t
83	30	68.2	222	1	SP12B VACCC	P20842 vaccinia vi
84	30	68.2	222	2	Q6J350 9PROXV	Q6J350 vaccinia vi
85	30	68.2	222	2	Q9UF39 VACCT	Q9UF39 vaccinia vi
86	30	68.2	224	2	Q9HP93 HAL9A	Q9HP93 halobacteri
87	30	68.2	224	2	Q9BBE9 PIG	Q9BBE9 sue scrofa
88	30	68.2	239	2	Q25653 PLABE	Q25653 plasmodium
89	30	68.2	240	2	Q59587 9PACTO	Q59587 streptomyce
90	30	68.2	246	2	Q5NNY6 ANOGA	Q5NNY6 anopheles g
91	30	68.2	263	2	Q58827 ECTHA	Q58827 ecotolhorno
92	30	68.2	290	2	Q4G0U2 HUMAN	Q4G0U2 homo sapien
93	30	68.2	299	2	Q5SM71 CRYNE	Q5SM71 cryptococcu
94	30	68.2	299	2	Q5KBS7 CRYNE	Q5KBS7 cryptococcu
95	30	68.2	329	2	Q6C1K4 YARLI	Q6C1K4 yarrowia li
96	30	68.2	336	2	Q9KDK4 BACHD	Q9KDK4 bacillus ha
97	30	68.2	345	1	SP12 RABPU	P42826 rabdipox v
98	30	68.2	345	1	SP12 VACCV	P15059 vaccinia vi
99	30	68.2	345	1	Q72749 COMXP	Q72749 composita v
100	30	68.2	347	2	Q4KSK4 9NUCL	Q4KSK4 chrysodeixi
101	30	68.2	359	2	Q4PLCQ SCOMA	Q4PLCQ schistosoma
102	30	68.2	397	2	Q7NKS5 GLOVI	Q7NKS5 gloeobacter
103	30	68.2	406	2	Q21712 CABEL	Q21712 caenorhabdi
104	30	68.2	423	2	Q4U1U6 9SMBG	Q4U1U6 punctatus p

105	30	68.2	426	2	041x25_9BURK	041x25_burkholderi	178	29	65.9	162	2	062139_STRAW	062139_streptomyce
106	30	68.2	430	2	021713_CAREL	021713_caenorhabdi	179	29	65.9	163	2	061832_CAREL	061832_caenorhabdi
107	30	68.2	442	2	084170_9POXY	084170_orf_virus.	180	29	65.9	164	2	061h39_CAREL	061h39_caenorhabdi
108	30	68.2	452	2	08TWY8_METKA	08twy8_methanopyru	181	29	65.9	171	2	06RP40_ARATH	06rf40_arabidopsis
109	30	68.2	452	2	06TWY1_9POXY	06twy1_orf_virus.	182	29	65.9	195	2	095025_CAREL	09sqz5_caenorhabdi
110	30	68.2	453	2	06TW91_9POXY	06tw91_orf_virus.	183	29	65.9	200	2	04fz56_PENMO	04fz56_pennaeus mon
111	30	68.2	479	2	07RYP8_NEUCR	07ryp8_neuosporea	184	29	65.9	201	2	061K16_DROME	061k16_drosophila
112	30	68.2	493	2	09ATG8_CHLRE	09atg8_chlamydomon	185	29	65.9	207	2	083NCG_TROM8	083nc8_tropheryma
113	30	68.2	527	1	MAS13_CAREL	020191_caenorhabdi	186	29	65.9	211	2	08VKC3_MYCTU	08vk33_mycobacteri
114	30	68.2	551	1	ARS_HENPU	P14000_hemicentrot	187	29	65.9	219	2	000048_AJDEE	000048_ajellomyces
115	30	68.2	551	1	Q25384_HEMPU	Q25384_hemicentrot	188	29	65.9	226	2	0673H2_PENVA	0673h2_pennaeus van
116	30	68.2	558	2	Q52GF0_MAGGR	Q52GF0_magnaporthe	189	29	65.9	226	2	04fz57_LITSE	04fz57_litopenaeus
117	30	68.2	572	1	CECR6_MOUSE	099x67_mus_musculu	190	29	65.9	226	2	07UL27_RHOBA	07ul27_rhodopiirell
118	30	68.2	574	2	075632_NEUCR	075632_neuosporea	191	29	65.9	228	2	Q4SRJ6_TETNG	Q4srj6_tetradodon n
119	30	68.2	582	2	Q6MPL4_NEUCR	Q6mfl4_neutospora	192	29	65.9	231	2	051KW9_MAGGR	051kw9_magnaporthe
120	30	68.2	619	2	Q4PJU7_USTMA	Q4pj7_ustiliago ma	193	29	65.9	237	2	054V93_DICDI	054v93_dictyosteli
121	30	68.2	635	2	Q912M7_RAT	Q912m7_rattus norv	194	29	65.9	245	2	Q9ZUV1_ARATH	Q9zuv1_arabidopsis
122	30	68.2	635	1	GBRT_MOUSE	0911f1_mus_musculu	195	29	65.9	254	2	0610G4_PICTO	0610g4_picrophilus
123	30	68.2	642	2	Q607E9_METCA	Q607e9_methylcococ	196	29	65.9	262	2	07NU98_GLOVI	07nj98_gloeobacter
124	30	68.2	650	2	Q4SSS2_TETNG	Q4SSS2_tetradodon n	197	29	65.9	263	2	051S12_MAGGR	051s12_magnaporthe
125	30	68.2	652	2	Q8NB41_HUMAN	Q8nb41_homo_sapien	198	29	65.9	271	2	Q6WUJ7_PENMO	Q6wuj7_pennaeus mon
126	30	68.2	699	2	Q8COW9_MOUSE	Q8c9w9_mus_musculu	199	29	65.9	274	2	Q7UI56_RHOBA	Q7ui56_rhodopiirell
127	30	68.2	707	2	Q8COW9_MOUSE	Q7yej4_dictyosteli	200	29	65.9	276	2	Q93X97_ANTHI	Q93x97_antirrhinum
128	30	68.2	715	2	Q94494_DICDI	Q94494_dictyosteli	201	29	65.9	276	2	Q4T6Q2_TETNG	Q4t6q2_tetradodon n
129	30	68.2	744	2	Q4T9V5_TETNG	Q4t9v5_tetradodon n	202	29	65.9	281	2	Q7UNK9_RHOBA	Q7unk9_rhodopiirell
130	30	68.2	746	2	Q54CH8_DICDI	Q54ch8_dictyosteli	203	29	65.9	287	2	05YOKO_ORYSA	05yqk0_oryza sativ
131	30	68.2	764	2	Q9HAP8_HUMAN	Q9hap8_homo_sapien	204	29	65.9	293	1	CHIE_BEYU	P36910_beta_vulgar
132	30	68.2	785	2	Q8TKM7_METAC	Q8tkm7_methanogarc	205	29	65.9	294	2	Q6F7P8_ACIAD	Q6f7p8_acetnecobact
133	30	68.2	789	2	Q9H8C9_HUMAN	Q9h8c9_homo_sapien	206	29	65.9	293	2	Q61H14_CAREL	Q61h14_caenorhabdi
134	30	68.2	834	2	Q5JHL1_PYRO	Q5jhl1_pyrococcus	207	29	65.9	326	2	Q74GE8_GEOSL	Q74ge8_geobacter b
135	30	68.2	849	2	Q5JHL1_PYRO	Q97776_eulemur ful	208	29	65.9	330	2	Q7M725_WOLSU	Q7m725_wolfinella s
136	30	68.2	884	1	ANDR_EDULC	Q9GK17_sus_scrofa	209	29	65.9	333	2	Q980J4_SULISO	Q980j4_sulfolobus
137	30	68.2	896	1	ANDR_PIG	Q54C31_dictyosteli	210	29	65.9	335	2	Q57UB4_GTRYP	Q57ub4_trypanosoma
138	30	68.2	966	2	Q54C31_DICDI	Q54c31_magnaporthe	211	29	65.9	336	2	Q6ZVW7_HUMAN	Q6zvw7_homo_sapien
139	30	68.2	1157	2	Q51UM6_MAGGR	Q51um6_magnaporthe	212	29	65.9	345	2	Q61G18_DROME	Q61g18_drosophila
140	30	68.2	1245	2	Q962K7_PLANT	Q962k7_plasmiodium	213	29	65.9	348	2	Q9X7M9_STROK	Q9x7m9_streptomyce
141	30	68.2	1367	2	Q4PHL3_USTMA	Q4phl3_ustiliago ma	214	29	65.9	350	2	Q60678_9BURK	Q60678_collimonas
142	30	68.2	1375	2	Q54YB9_DICDI	Q54yb9_dictyosteli	215	29	65.9	350	2	Q82MH4_STRAW	Q82mh4_streptomyce
143	30	68.2	1556	2	Q42Z73_PLABE	Q42z73_plasmiodium	216	29	65.9	352	2	Q5FW56_MOUSE	Q5fw56_mus_musculu
144	30	68.2	1721	2	Q614N6_CABER	Q614n6_caenorhabdi	217	29	65.9	351	2	Q50005_PERSY	Q50005_pseudomonas
145	30	68.2	1722	2	Q19350_CAREL	Q19350_caenorhabdi	218	29	65.9	351	2	Q88B73_PERSM	Q88b73_pseudomonas
146	30	68.2	1791	2	Q25645_PLABE	Q25645_plasmiodium	219	29	65.9	351	2	Q8VFM3_9MICC	Q8vfm3_micrococcus
147	30	68.2	1791	2	Q906D4_PLABE	Q906d4_leishmania	220	29	65.9	356	2	Q4FVU6_LEIMA	Q4fvu6_leishmania
148	30	68.2	1905	2	Q4Q6D0_LEIMA	Q4q6d0_leishmania	221	29	65.9	357	2	Q52B09_MAGGR	Q52b09_magnaporthe
149	30	68.2	1972	2	Q7PUY1_ANGOA	Q7pyu1_anopheles g	222	29	65.9	366	2	Q41ZRS_AZOVI	Q41zrs_azotobacter
150	30	68.2	2161	2	Q91301_9VIRU	Q91301_rupestrilis s	223	29	65.9	369	2	Q4KKM4_PSEF5	Q4kkm4_pseudomonas
151	30	68.2	2161	2	Q52W13_9VIRU	Q52w13_rupestrilis s	224	29	65.9	375	2	Q91766_PSEAB	Q91766_pseudomonas
152	30	68.2	2161	2	Q52W18_9VIRU	Q52w18_rupestrilis s	225	29	65.9	379	2	Q04326_PICAN	Q04326_pichia angu
153	30	68.2	23015	2	Q81O18_DROME	Q81q18_drosophila	226	29	65.9	383	2	Q5P3Z8_AZOSE	Q5p3z8_azarcus sp
154	29	65.9	45	2	Q91W92_TOBAC	Q91w92_nicotiana t	227	29	65.9	385	2	Q8DLR1_SYNEK	Q8dlr1_synechococc
155	29	65.9	46	1	Y940_TEPBA	Q83910_treponema p	228	29	65.9	387	2	Q88RN6_PSEPK	Q88rn6_pseudomonas
156	29	65.9	76	2	Q4SYE8_TETNG	Q4sy88_tetradodon n	229	29	65.9	394	1	K1H4_HUMAN	Q76011_homo_sapien
157	29	65.9	94	2	Q75079_HUMAN	Q4sy88_tetradodon n	230	29	65.9	395	2	Q8G7M2_BIFLO	Q8g7m2_bifidobacte
158	29	65.9	95	2	Q5TA77_HUMAN	Q5ta77_homo_sapien	231	29	65.9	404	2	Q9PULO_BRABE	Q9pulo_brechydantio
159	29	65.9	116	2	Q6R647_CAPHI	Q6r647_capra hircu	232	29	65.9	407	2	Q5B3N5_EMENI	Q5b3n5_aspergillus
160	29	65.9	118	2	Q53NF3_ORYSA	Q53nf3_oryza sativ	233	29	65.9	415	2	Q5KC03_CRYNE	Q5kc03_cryptococcu
161	29	65.9	119	1	CHCA1_BOOMO	P13531_bombyx mori	234	29	65.9	429	2	Q8MRX5_DROME	Q8mrx5_drosophila
162	29	65.9	119	1	CHCA2_BOOMO	Q17212_bombyx mori	235	29	65.9	435	2	Q9VNP4_DROME	Q9vnp4_drosophila
163	29	65.9	119	1	CHCA3_BOOMO	Q17213_bombyx mori	236	29	65.9	435	2	Q4S0J5_TETNG	Q4s0j5_tetradodon n
164	29	65.9	121	1	CHCA5_BOOMO	Q17214_bombyx mori	237	29	65.9	436	2	Q81UT8_HUMAN	Q81ut8_homo_sapien
165	29	65.9	123	2	Q5N7A4_ORYSA	Q5n7a4_oryza sativ	238	29	65.9	444	2	Q51KF5_MAGGR	Q51kf5_magnaporthe
166	29	65.9	126	2	Q58HG2_9MXI	Q5n7a4_oryza sativ	239	29	65.9	449	2	Q9S9K4_ARATH	Q9s9k4_arabidopsis
167	29	65.9	129	2	Q6ZSH8_HUMAN	Q6zsh8_polyascaris p	240	29	65.9	464	2	Q86GK1_GTRYP	Q86gk1_trypanosoma
168	29	65.9	129	2	Q6ZSH8_HUMAN	Q6zsh8_homo_sapien	241	29	65.9	469	2	Q8SWY9_DROME	Q8swy9_drosophila
169	29	65.9	131	2	Q5OLJ9_ORYSA	Q5olj9_oryza sativ	242	29	65.9	472	2	Q6Q466_9TRYP	Q6q466_trypanosoma
170	29	65.9	133	2	Q8MOJ7_CAREL	Q8mqj7_caenorhabdi	243	29	65.9	475	2	Q4Q4Q0_LEIMA	Q4q4q0_leishmania
171	29	65.9	147	2	Q9PYH8_XYLEPA	Q9pyh8_xylella fas	244	29	65.9	475	2	Q8H3W6_ORYSA	Q8h3w6_oryza sativ
172	29	65.9	151	2	Q4OCK6_LEIMA	Q4qck6_leishmania	245	29	65.9	481	2	Q968M0_9TRYP	Q968m0_trypanosoma
173	29	65.9	151	2	Q7YS88_PIG	Q7ys88_sus_scrofa	246	29	65.9	481	2	Q58DB9_BOVIN	Q58db9_bos taurus
174	29	65.9	151	2	Q867B8_PIG	Q867b8_sus_scrofa	247	29	65.9	489	2	Q5WRM0_CAREL	Q5wrm0_magnaporthe
175	29	65.9	153	2	Q5U614_HUMAN	Q5u614_homo_sapien	248	29	65.9	496	2	Q9RG51_BACCI	Q9rg51_bacillus ci
176	29	65.9	153	2	Q5U614_HUMAN	Q5u614_homo_sapien	249	29	65.9	504	2	Q54PDB_DICDI	Q54pdb_dictyosteli
177	29	65.9	161	2	Q61HS7_CABER	Q61hs7_caenorhabdi	250	29	65.9	504	2	Q54PDB_DICDI	Q54pdb_dictyosteli



251	29	65.9	506	2	Q60QH2_CABER	Q60qh2 caenorhabdi	324	28	63.6	74	1	MS94B_DROME	Q01643 drosophila
252	29	65.9	507	2	Q9XUV9_CABEL	Q9Xuv9 caenorhabdi	325	28	63.6	79	2	Q6K3AI_ORYSA	Q6K3ai oryza sativ
253	29	65.9	515	1	APX1_CABEL	P41990 caenorhabdi	326	28	63.6	74	2	Q96UL9_AJECA	Q96ul9 ajeelomyces
254	29	65.9	515	2	Q4T2J6_TERNG	Q4c7j6 tetraodon n	327	28	63.6	86	2	Q8UMD7_9H1V1	Q8umd7 human immun
255	29	65.9	541	1	DCR1B_MOUSE	Q8c7j6 mus musculu	328	28	63.6	91	2	Q9NCR3_9CUCU	Q9ncr3 drosophila
256	29	65.9	547	2	Q7UVV3_RHOBA	Q7uvv3 rhodopirell	329	28	63.6	91	2	Q18928_MACRA	Q18928 macaca radi
257	29	65.9	551	2	Q9DIC6_9RHAB	Q9dic6 lettuce nec	330	28	63.6	95	2	Q8URK1_AGR75	Q8urk1 agrobacteri
258	29	65.9	574	2	Q9BMT7_DROME	Q9bmt7 drosophila	331	28	63.6	96	2	Q96419_9CUCU	Q96419 drosophila
259	29	65.9	574	2	Q9VJ22_DROME	Q9vj22 drosophila	332	28	63.6	96	2	Q5LAC3_BACFN	Q5lac3 bacteroides
260	29	65.9	574	2	Q6GWC3_BRAE	Q6gwc3 brachydantio	333	28	63.6	98	2	Q9NCR6_9CUCU	Q9ncr6 drosophila
261	29	65.9	575	2	Q7QNV8_GIALA	Q7qnv8 giardia lam	334	28	63.6	101	2	Q4QXB8_9H1V1	Q4qxb8 human immun
262	29	65.9	575	2	Q57U16_9TRYP	Q57u16 trypanosoma	335	28	63.6	103	2	Q49320_ARYTH	Q49320 arabidopsis
263	29	65.9	578	2	Q4T917_TERNG	Q4t917 tetraodon n	336	28	63.6	105	2	Q17187_BOMBO	Q17187 bombyx mori
264	29	65.9	612	2	Q9KZU9_STRCO	Q9kzu9 streptomyces	337	28	63.6	106	2	Q9GSM7_9HYME	Q9gsm7 xylocopa ol
265	29	65.9	614	2	Q4RLCS_TERNG	Q4rlcs tetraodon n	338	28	63.6	107	2	Q8RKM5_PPROE	Q8rkm5 providencia
266	29	65.9	630	2	Q6TNG4_PICPA	Q6tng4 picchia fari	339	28	63.6	108	2	Q46351_9CUCU	Q46351 drosophila
267	29	65.9	680	2	Q72WMS_DESVH	Q72wms desulfovibr	340	28	63.6	108	2	Q46352_9CUCU	Q46352 drosophila
268	29	65.9	685	1	DL4_HUMAN	Q9nrf61 homo sapien	341	28	63.6	108	2	Q5LAC4_BACFN	Q5lac4 bacteroides
269	29	65.9	686	1	DL4_MOUSE	Q9j171 mus musculu	342	28	63.6	108	2	Q64QD5_BACFR	Q64qd5 bacteroides
270	29	65.9	686	2	Q9DBU9_MOUSE	Q9dbu9 mus musculu	343	28	63.6	109	2	Q46346_9CUCU	Q46346 drosophila
271	29	65.9	728	2	Q54DV5_DICDI	Q54dv5 dictyosteli	344	28	63.6	110	2	Q15739_HUMAN	Q15739 homo sapien
272	29	65.9	733	2	Q54P69_DICDI	Q54p69 dictyosteli	345	28	63.6	110	2	Q15739_HUMAN	Q15739 homo sapien
273	29	65.9	748	2	Q86HR8_DICDI	Q86hr8 dictyosteli	346	28	63.6	115	1	MERT_SHEPU	Q54462 shewanella
274	29	65.9	764	2	Q97343_SUDDA	Q97343 suberites d	347	28	63.6	120	2	Q86BW2_ASCSS	Q86bw2 ascidia syd
275	29	65.9	830	2	Q6INMO_XENILA	Q6inmo xenopus lae	348	28	63.6	121	2	Q9NCR1_9CUCU	Q9ncr1 dendroideles
276	29	65.9	843	1	CO7_HUMAN	P10643 homo sapien	349	28	63.6	121	2	Q9RMS0_BRAJA	Q9rms0 brachydantio
277	29	65.9	843	1	CO7_PIG	Q9rtq3 sus scrofa	350	28	63.6	122	2	Q9NCR9_9CUCU	Q9ncr9 dendroideles
278	29	65.9	843	2	Q5RADO_PONPY	Q5rado pongo pygma	351	28	63.6	123	2	Q9NCR9_9CUCU	Q9ncr9 dendroideles
279	29	65.9	919	2	Q61V24_CABER	Q61v24 caenorhabdi	352	28	63.6	129	2	Q9NCR9_9CUCU	Q9ncr9 dendroideles
280	29	65.9	934	2	Q7XVG0_ORYSA	Q7xvg0 oryza sativ	353	28	63.6	131	2	Q4T174_TERNG	Q4t174 tetraodon n
281	29	65.9	966	2	Q22378_CABEL	Q22378 caenorhabdi	354	28	63.6	132	2	Q7XJVO_ORYSA	Q7xjvo oryza sativ
282	29	65.9	1005	2	Q5ZM51_CHICK	Q5zm51 gallus gall	355	28	63.6	136	2	Q9GS67_9HYME	Q9gs67 xylocopa tr
283	29	65.9	1029	2	Q9VW85_DROME	Q9vw85 drosophila	356	28	63.6	139	2	Q9GS73_9HYME	Q9gs73 xylocopa tr
284	29	65.9	1053	2	Q69ZC3_MOUSE	Q69zc3 mus musculu	357	28	63.6	141	2	Q8AW84_BRAE	Q8aw84 brachydantio
285	29	65.9	1070	2	Q5AKJ3_EMENT	Q5akj3 aspergillus	358	28	63.6	143	2	Q9GS66_XYLV1	Q9gs66 xylocopa vi
286	29	65.9	1085	2	Q4SEL7_TERNG	Q4sel7 tetraodon n	359	28	63.6	144	2	Q5N2D3_SYNP6	Q5n2d3 synechococ
287	29	65.9	1094	2	Q7SD86_NEUCR	Q7sd86 neurospora	360	28	63.6	146	2	Q4WGI8_ASEPT	Q4wgi8 aspergillus
288	29	65.9	1106	2	P79726_BRAE	P79726 brachydantio	361	28	63.6	146	2	Q5VT82_HUMAN	Q5vt82 homo sapien
289	29	65.9	1142	2	Q615H8_ORYSA	Q615h8 oryza sativ	362	28	63.6	146	2	Q4YMU1_PLABE	Q4ymu1 plasmodium
290	29	65.9	1183	2	Q7SDP1_NEUCR	Q7sdp1 neurospora	363	28	63.6	146	2	Q8BSG4_MOUSE	Q8bsg4 mus musculu
291	29	65.9	1224	2	Q4SEB5_TERNG	Q4seb5 tetraodon n	364	28	63.6	147	2	Q9GS71_9HYME	Q9gs71 xylocopa vi
292	29	65.9	1335	2	Q9FME7_ARYTH	Q9fme7 arabidopsis	365	28	63.6	147	2	Q5TVP3_ANOGA	Q5tvp3 anopheles g
293	29	65.9	1357	2	Q5MD89_BRAE	Q5md89 brachydantio	366	28	63.6	149	2	Q8AVJ9_BRAE	Q8avj9 brachydantio
294	29	65.9	1357	2	Q5GIT2_BRAE	Q5git2 brachydantio	367	28	63.6	152	2	Q8FYU8_BRURU	Q8fyu8 bruceella su
295	29	65.9	1387	2	Q4SLQ3_TERNG	Q4slq3 tetraodon n	368	28	63.6	153	2	Q9GS64_9HYME	Q9gs64 xylocopa su
296	29	65.9	1417	2	Q4QCG2_LEIMA	Q4qcg2 leishmania	369	28	63.6	153	2	Q9GS70_9HYME	Q9gs70 xylocopa bo
297	29	65.9	1420	2	Q4SCY1_TERNG	Q4scy1 tetraodon n	370	28	63.6	153	2	Q9GS72_9HYME	Q9gs72 xylocopa au
298	29	65.9	1423	1	Q5SW75_MOUSE	Q5sw75 mus musculu	371	28	63.6	154	2	Q9YAZ4_AERYE	Q9yaz4 aetopyrum p
299	29	65.9	1427	1	MS84_DROME	Q8m36 drosophila	372	28	63.6	156	2	Q6Y262_9EURY	Q6y262 uncultured
300	29	65.9	1564	2	Q86MD7_ECHMU	Q86md7 echinococcu	373	28	63.6	158	2	Q7MT99_PORDI	Q7mt99 pophyromon
301	29	65.9	1637	2	Q9XSIV8_BOVIN	Q9xsv8 bos taurus	374	28	63.6	162	2	Q9BLH6_APLXU	Q9blh6 aplysia xur
302	29	65.9	1743	2	Q9XWXS_CABEL	Q9xwx5 caenorhabdi	375	28	63.6	163	2	Q4SEL6_TERNG	Q4sel6 tetraodon n
303	29	65.9	1748	2	Q61TPI_CABER	Q61tpi caenorhabdi	376	28	63.6	164	2	Q5DY73_VIBP1	Q5dy73 vibrio fusc
304	29	65.9	1755	2	Q55BY2_DICDI	Q55by2 dictyosteli	377	28	63.6	165	2	Q26742_9TRYP	Q26742 trypanosoma
305	29	65.9	1843	2	Q9GQ82_LEITR	Q9gq82 leishmania	378	28	63.6	168	2	Q5NXP6_AZORSE	Q5npx6 azaracus sp
306	29	65.9	1917	2	Q51SM3_ENTHI	Q51sw3 entamoeba h	379	28	63.6	173	2	Q51F05_HORSE	Q51f05 equus cabal
307	29	65.9	2117	2	Q4QIS6_LEIMA	Q4qis6 leishmania	380	28	63.6	176	2	Q879W2_XYLET	Q879w2 xyella fas
308	29	65.9	2146	2	Q4T7A2_TERNG	Q4t7a2 tetraodon n	381	28	63.6	176	2	Q891S1_BRAJA	Q891s1 bradyrhizob
309	29	65.9	2183	2	Q64962_9VIRU	Q64962 apple stem	382	28	63.6	176	2	Q6J340_9PROX	Q6j340 vaccinia vi
310	29	65.9	2185	2	Q9E948_9VIRU	Q9e948 apple stem	383	28	63.6	177	2	Q6AS16_ORYSA	Q6as16 oryza sativ
311	29	65.9	2938	2	Q41R68_GIBZE	Q41r68 gibberella	384	28	63.6	178	2	P74672_SYNY3	P74672 synechocyst
312	29	65.9	3434	2	Q4G444_XENILA	Q4g444 xenopus lae	385	28	63.6	194	2	Q90YVC3_CHICK	Q90yvc3 gallus gall
313	29	65.9	3695	1	LAMAS_HUMAN	Q15130 homo sapien	386	28	63.6	196	1	ZCSL3_MOUSE	Q19120 mus musculu
314	29	65.9	3695	2	Q8TDF6_HUMAN	Q8tdf6 homo sapien	387	28	63.6	198	2	Q9RCU7_STRCO	Q9rcu7 streptomyces
315	29	65.9	3718	1	LAMAS_MOUSE	Q61001 mus musculu	388	28	63.6	201	2	Q9ANB9_BRAJA	Q9anb9 bradyrhizob
316	29	65.9	4998	1	Q8CG65_MOUSE	Q8cg65 mus musculu	389	28	63.6	202	2	Q8H717_ORYSA	Q8h717 oryza sativ
317	29	65.9	5141	2	Q700K0_RAP	Q700k0 rattus norv	390	28	63.6	202	2	Q8S796_ORYSA	Q8s796 oryza sativ
318	29	65.9	5146	2	Q8SPM4_BOVIN	Q8spm4 bos taurus	391	28	63.6	203	2	Q27654_9NEOP	Q27654 extoxocrania
319	28	63.6	25	1	LRP3_CONPU	P60245 conus purpu	392	28	63.6	207	1	FLAB2_ARCFU	Q29207 magnaesglobe
320	28	63.6	25	1	LRP3_CONPU	P83629 felis silve	393	28	63.6	207	2	Q52G10_MAGGR	Q52g10 magnaporthe
321	28	63.6	31	2	Q4YAB1_PLABE	Q4yab1 plasmodium	394	28	63.6	207	2	Q9ZM78_HELPU	Q9zm78 helicobacte
322	28	63.6	55	2	Q7UXB5_RHOBA	Q7uxb5 rhodopirell	395	28	63.6	209	2	P64146_TRILE	P64146 trichosanth
323	28	63.6	59	2	Q6ADE4_LEITX	Q6ade4 leifsonia x	396	28	63.6	210	2	Q7X9J2_WHEAT	Q7x9j2 triticum ae

397	28	63.6	215	2	051P06_MAGGR	051p06 magnaporthe	470	28	63.6	450	1	LIPP_PIG	P00591 sus scrofa
398	28	63.6	215	2	04S2L6_TERNG	04s2l6 terradon n	471	28	63.6	450	2	06CAT3_VARLT	06cat3 yarrowia li
399	28	63.6	218	2	09T7F3_CABEL	09t7f3 caenorhabdi	472	28	63.6	462	2	08MOC6_CABEL	08moc6 caenorhabdi
400	28	63.6	234	2	07G3Y2_ANOGA	07g3y2 anopheles g	473	28	63.6	464	1	08NFR1_MOUSE	08nfr1 mus musculus
401	28	63.6	237	1	1BP4_SHEEP	028893 ovis aries	474	28	63.6	465	2	07XIX6_ORYSA	07xix6 oryza sativ
402	28	63.6	239	2	054K45_DICDI	054k45 dictyostell	475	28	63.6	471	2	05U3U4_BRARE	05u3u4 brachydanio
403	28	63.6	248	2	08X6T7_ECOS7	08x6t7 escherichia	476	28	63.6	474	2	07Y0G5_ORYSA	07y0g5 oryza sativ
404	28	63.6	254	1	1BP4_MOUSE	047879 mus musculus	477	28	63.6	485	2	07ULG9_RHOBA	07ulg9 rhodospirill
405	28	63.6	254	1	1BP4_RAT	P21744 rattus norv	478	28	63.6	491	2	09A7T1_CAUCR	09a7t1 caulobacter
406	28	63.6	254	1	0946U5_HORVU	0946u5 hordelium vul	479	28	63.6	493	2	051P74_MAGGR	051p74 magnaporthe
407	28	63.6	254	2	08BSM9_MOUSE	08bsm9 mus musculus	480	28	63.6	507	2	08R0K2_MOUSE	08r0k2 mus musculus
408	28	63.6	254	2	06B374_RAT	06b374 rattus norv	481	28	63.6	508	2	07S178_ORYSA	07s178 oryza sativ
409	28	63.6	258	1	1BP4_BOVIN	005716 bos taurus	482	28	63.6	509	2	07Q0M5_ANOGA	07q0m5 anopheles g
410	28	63.6	258	1	1BP4_HUMAN	P22692 homo sapien	483	28	63.6	521	2	09BAL0_RHIL0	09bal0 rhizobium l
411	28	63.6	258	2	05U0I2_HUMAN	05u0i2 homo sapien	484	28	63.6	526	2	04FKN6_TRRYP	04fkn6 trypanosoma
412	28	63.6	268	2	09W1S2_DROME	09w1s2 drosophila	485	28	63.6	531	2	09BZG5_HUMAN	09bzg5 homo sapien
413	28	63.6	270	2	054201_DICDI	054201 dictyostell	486	28	63.6	535	2	07R3T0_GITALA	07r3t0 giardia lam
414	28	63.6	272	2	069079_HHVI	069079 human herpe	487	28	63.6	536	2	04NNP1_9BELT	04nnp1 anaeromyxob
415	28	63.6	276	2	04Q092_DROME	04q092 drosophila	488	28	63.6	538	2	05USC7_HUMAN	05usc7 homo sapien
416	28	63.6	276	2	08LJP5_GLAMI	08ljp5 antirrhinum	489	28	63.6	539	2	09NUA2_HUMAN	09nu2 homo sapien
417	28	63.6	276	2	04TTH5_TERNG	04th5 terradon n	490	28	63.6	542	2	09BZG6_HUMAN	09bzg6 homo sapien
418	28	63.6	278	2	06K1W1_ORYSA	06k1w1 oryza sativ	491	28	63.6	544	2	09BZG7_HUMAN	09bzg7 homo sapien
419	28	63.6	282	2	082RB2_STRAW	082rb2 streptomyce	492	28	63.6	544	2	056M91_ARATH	056m91 arabidopsis
420	28	63.6	284	2	054B88_DICDI	054b88 dictyostell	493	28	63.6	552	2	089RV8_BRADJ	089rv8 bradyrhizob
421	28	63.6	288	2	001848_PNECA	001848 pneumocysti	494	28	63.6	557	2	0968Z6_DICDI	0968z6 dictyostell
422	28	63.6	293	2	07PMK7_ANOGA	07pmk7 anopheles g	495	28	63.6	559	2	09VZ44_DROME	09vz44 drosophila
423	28	63.6	294	2	09RCF3_STROCA	09rcf3 streptomyce	496	28	63.6	560	2	054GF7_DICDI	054gf7 dictyostell
424	28	63.6	296	2	09UUF3_PNECA	09uuf3 pneumocysti	497	28	63.6	562	2	04RHZ7_TERNG	04rh27 terradon n
425	28	63.6	299	2	04Q1A5_LEIMA	04q1a5 leishmania	498	28	63.6	567	1	ARS_STRPU	P50473 streptomycoc
426	28	63.6	305	2	09YFP7_HHVI	09yfp7 human herpe	499	28	63.6	569	2	09RFF2_RHOSH	09rff2 rhodobacter
427	28	63.6	311	2	0580R2_9TRYP	0580r2 trypanosoma	500	28	63.6	572	2	06NB73_RHOPA	06nb73 rhodospheuo
428	28	63.6	315	2	08KR56_ESCPE	08kr56 escherichia	501	28	63.6	575	2	081RB8_DROME	081rb8 drosophila
429	28	63.6	315	2	0910J4_STROCA	0910j4 streptomyce	502	28	63.6	588	2	0580V3_TRRYP	0580v3 trypanosoma
430	28	63.6	316	2	04S672_TERNG	04s672 terradon n	503	28	63.6	591	2	073P17_TREDE	073p17 treponema d
431	28	63.6	316	2	08ZHB3_YERPE	08zhb3 yerquina pe	504	28	63.6	594	2	084Z97_ORYSA	084z97 oryza sativ
432	28	63.6	316	2	0666J2_YERPS	0666j2 yerquina ps	505	28	63.6	599	2	0833X0_ENTFA	0833x0 enterococcu
433	28	63.6	327	2	086C02_VIBVU	086c02 vibrio vuln	506	28	63.6	599	2	04QB90_LEIMA	04qb90 leishmania
434	28	63.6	328	2	04RP65_TERNG	04rp65 terradon n	507	28	63.6	603	2	08ZG41_STRAW	08zg41 streptomyce
435	28	63.6	331	2	057KM1_SALCH	057km1 salmonella	508	28	63.6	617	2	054NF7_DICDI	054nf7 dictyostell
436	28	63.6	331	2	05PFO0_SALPA	05pfo0 salmonella	509	28	63.6	619	2	09YMI9_DROME	09ym19 drosophila
437	28	63.6	331	2	08Z4E2_SALTY	08z4e2 salmonella	510	28	63.6	624	1	SUYH4_ARATH	084297 arabidopsis
438	28	63.6	331	2	08ZML1_SALTY	08zml1 salmonella	511	28	63.6	628	1	STTL_YEAST	P39980 saccharomyc
439	28	63.6	334	2	09NZS8_HUMAN	09nzs8 homo sapien	512	28	63.6	628	2	0619B9_CABER	0619c9 caenorhabdi
440	28	63.6	334	2	06DIE4_ERMCT	06die4 erwina car	513	28	63.6	632	1	GBRT_HUMAN	09un88 homo sapien
441	28	63.6	336	2	04J3I4_AZOVI	04j3i4 azorobacter	514	28	63.6	632	2	06WGM0_9PARA	06wgm0 mossman vir
442	28	63.6	340	1	LRP1_HHVI	P17888 human herpe	515	28	63.6	632	2	06CX82_KLUTA	06cx82 kluyveromyc
443	28	63.6	354	2	061AK0_CABER	061ak0 caenorhabdi	516	28	63.6	634	2	06CX82_KLUTA	06cx82 kluyveromyc
444	28	63.6	355	2	07S6V6_NEUCR	07s6v6 neurospora	517	28	63.6	648	2	04S516_TERNG	04s516 tetradon n
445	28	63.6	360	2	09H6B9_HUMAN	09h6b9 homo sapien	518	28	63.6	654	2	04Q5A1_LEIMA	04q5a1 leishmania
446	28	63.6	360	2	08ECB2_SHEON	08ecb2 shevanella	519	28	63.6	655	2	06BV47_DBBHA	06bv47 debaryomyce
447	28	63.6	364	2	07YMW5_CABEL	07ymw5 caenorhabdi	520	28	63.6	663	2	09VTX2_DROME	09vtx2 drosophila
448	28	63.6	368	2	04VNA3_MOUSE	04vna3 mus musculus	521	28	63.6	668	2	05D734_BRARE	05d734 brachydanio
449	28	63.6	375	2	0946G9_PARP	0946g9 parametium	522	28	63.6	672	2	04IK90_GIBZE	04ik90 gibberella
450	28	63.6	375	2	04RMC1_TERNG	04rmc1 tetradon n	523	28	63.6	679	2	08PCTT_XANAC	08pctt xanthomonas
451	28	63.6	379	2	08Z3I4_CHLCV	08z3i4 chlamydomophi	524	28	63.6	696	2	053RS8_HUMAN	053rs8 homo sapien
452	28	63.6	381	2	09N5Y3_CABEL	09n5y3 caenorhabdi	525	28	63.6	696	2	09H8R3_HUMAN	09h8r3 homo sapien
453	28	63.6	383	2	06Z543_GEOCY	06z543 geodia cydo	526	28	63.6	700	2	06TVG0_9POXV	06tvgo bovine papu
454	28	63.6	384	2	05Z5R8_MAGGR	05z5r8 magnaporthe	527	28	63.6	703	2	05S1X6_CRYNE	05s1x6 cryptococcu
455	28	63.6	385	2	075GJ0_ORYSA	075gjo oryza sativ	528	28	63.6	703	2	05KCS7_CRYNE	05kcs7 cryptococcu
456	28	63.6	401	2	06LSH6_ORYSA	06lsh6 oryza sativ	529	28	63.6	705	2	04N0S9_9BELT	04n0s9 anaeromyxob
457	28	63.6	401	2	06NZ11_BRARE	06nzi1 brachydanio	530	28	63.6	706	2	086H21_DICDI	086h21 dictyostell
458	28	63.6	411	2	06OMF9_CABER	06omf9 caenorhabdi	531	28	63.6	709	2	06TVU2_9POXV	06tvu2 orf virus
459	28	63.6	412	2	04J3I5_AZOVI	04j3i5 azorobacter	532	28	63.6	716	2	09UB84_9CIL1	09ub84 parametium
460	28	63.6	420	2	06FK82_CANGA	06fk82 candida gla	533	28	63.6	720	2	06TW72_9POXV	06tw72 orf virus
461	28	63.6	428	2	0786D3_MAGGR	0786d3 magnaporthe	534	28	63.6	720	2	06MM20_BDBBA	06mm20 bdellovibri
462	28	63.6	428	2	0786D3_MAGGR	0786d3 magnaporthe	535	28	63.6	721	2	04L191_DICLA	04l191 dicentrarch
463	28	63.6	435	2	0616G8_CABER	0616g8 caenorhabdi	536	28	63.6	745	2	081YA6_HUMAN	081ya6 homo sapien
464	28	63.6	437	2	0960H1_HUMAN	0960h1 homo sapien	537	28	63.6	750	2	054VS4_DICDI	054vs4 dictyostell
465	28	63.6	437	2	08HDA5_HUMAN	08hda5 homo sapien	538	28	63.6	756	2	054VS2_DICDI	054vs2 dictyostell
466	28	63.6	437	2	08N6O3_HUMAN	08n6o3 homo sapien	539	28	63.6	786	2	04L1J2_GIBZE	04l1j2 gibberella
467	28	63.6	444	1	OPCA_STYMP	054709 synchococc	540	28	63.6	801	2	05G7N6_9CAUD	05g7n6 listonella
468	28	63.6	445	2	05N1G1_SYNP6	05n1g1 synchococc	541	28	63.6	803	1	C13AA_BACTU	045735 bacillus th
469	28	63.6	447	2	054VM4_DICDI	054vm4 dictyostell	542	28	63.6	809	2	05S2D5_CRYNE	05s2d5 cryptococcu

543	28	63.6	809	2	Q5KNQ1_CRYNE	Q5knq1_cryptococcu	616	28	63.6	2319	2	Q08H614_MAIZE	Q08h614_zea mays (m
544	28	63.6	833	2	Q6J288_ARACH	Q6j288_acanthamoeb	617	28	63.6	2395	2	Q27167_PARTE	Q27167_parametium
545	28	63.6	873	2	Q9LP01_ARATH	Q9lp01_arabidopsis	618	28	63.6	2597	2	Q6Q7Y4_PAPR	Q6q7y4_parametium
546	28	63.6	877	2	Q8WZU1_ORYSA	Q8wzu1_oryza sativ	619	28	63.6	2533	2	Q90589_PARTE	Q90589_parametium
547	28	63.6	895	1	ANDR_MACPA	Q97552_maccaca faec	620	28	63.6	2533	2	Q27183_PARTE	Q27183_parametium
548	28	63.6	895	1	ANDR_MACCU	Q6q755_maccaca mula	621	28	63.6	2543	2	P90649_PAPR	P90649_parametium
549	28	63.6	895	1	ANDR_PAPHA	Q97960_papio hamad	622	28	63.6	2551	2	STAB2_HUMAN	Q08wqz_h stablilin-
550	28	63.6	906	2	Q9UN21_HUMAN	Q9un21_homo sapien	623	28	63.6	2704	1	G168_PAPR	P17053_parametium
551	28	63.6	911	1	ANDR_PANTR	Q97775_pan troglod	624	28	63.6	2715	1	G156_PAPR	P13837_parametium
552	28	63.6	919	1	ANDR_HUMAN	P10275_homo sapien	625	28	63.6	2717	2	Q94710_PARTE	Q94710_parametium
553	28	63.6	920	2	Q5UTN9_HUMAN	Q5jtn9_homo sapien	626	28	63.6	2729	2	Q6PQK6_PARTE	Q6pqk6_parametium
554	28	63.6	930	2	Q9N777_HUMAN	Q9n777_homo sapien	627	28	63.6	3443	2	Q8U2M6_MOUSE	Q8u2m6_mus musculu
555	28	63.6	930	2	Q7RSC2_GIALA	Q7rsc2_giardia lam	628	27	61.4	34	1	TX1_STRCP	P60991_stromatopel
556	28	63.6	966	2	Q7UXM4_RHOBA	Q7uxm4_rhodopirell	629	27	61.4	36	1	TUT1A_HADVE	P82227_hadronyche
557	28	63.6	975	2	Q5WRLO_CABEL	Q5wrl0_caenorhabdi	630	27	61.4	36	1	TUT1B_HADVE	P82227_hadronyche
558	28	63.6	980	2	Q24780_CLOTM	Q24780_clostridial	631	27	61.4	37	1	TUT1C_HADVE	P82228_hadronyche
559	28	63.6	1000	2	Q4FKF7_9TRYX	Q4fkf7_trypanosoma	632	27	61.4	38	2	Q6VYR3D_THYCC	Q6vyr3d_thadronyche
560	28	63.6	1027	2	Q6FWT4_CANGA	Q6fwt4_candida gla	633	27	61.4	41	1	TX482_HYSGI	P56854_hycteroctat
561	28	63.6	1035	1	CDC68_YEAST	P32558_saccharomyc	634	27	61.4	43	2	Q6VQF7_CRAVI	Q6vqp7_cirsosostrea
562	28	63.6	1051	2	Q5U4U1_XENILA	Q5u4u1_xenopus lae	635	27	61.4	47	2	CRAM_CRAAB	P01542_cirambe abys
563	28	63.6	1084	2	Q7SM18_ORYSA	Q7sm18_oryza sativ	636	27	61.4	54	2	Q9N9H0_9BIVA	Q9n9h0_venerupis p
564	28	63.6	1088	1	MMT1_HORVU	Q9mbc2_hordeum vul	637	27	61.4	54	2	Q967H8_9BIVA	Q967h8_cirsosostrea
565	28	63.6	1091	1	MMT1_MAIZE	Q8w519_zea mays (m	638	27	61.4	55	2	Q7EZU4_ORYSA	Q7ezj4_oryza sativ
566	28	63.6	1115	2	Q4UH06_THRAN	Q4uh06_thelateria a	639	27	61.4	55	2	Q7LZAS_9PIPI	Q7lzas_xenopus sp.
567	28	63.6	1125	2	Q9N775_HUMAN	Q9n775_homo sapien	640	27	61.4	57	2	Q9N9H2_9BIVA	Q9n9h2_venerupis (
568	28	63.6	1128	2	Q4S6G8_TETNG	Q4s6g8_tetradodon n	641	27	61.4	64	2	Q9N9H1_9BIVA	Q9n9h1_venerupis (
569	28	63.6	1139	2	Q54Z39_DICDI	Q54z39_dicyostell	642	27	61.4	64	2	Q9PH34_XYLEFA	Q9ph34_xylella fas
570	28	63.6	1146	2	Q4R8T6_MACPA	Q4r8t6_maccaca faec	643	27	61.4	69	2	Q64VGS_ARATH	Q64vgs_arabidopsis
571	28	63.6	1156	2	Q9S938_HUMAN	Q9s938_homo sapien	644	27	61.4	71	2	Q8WQ16_OSTED	Q8wq16_ostrea edul
572	28	63.6	1160	2	Q4RX89_TETNG	Q4rx89_tetradodon n	645	27	61.4	71	2	Q80K29_9PAPI	Q80k29_human papil
573	28	63.6	1172	2	Q9LP05_ARATH	Q9lp05_arabidopsis	646	27	61.4	72	1	CXAS2_CONST	P28879_conus stria
574	28	63.6	1176	2	Q9N776_HUMAN	Q9n776_homo sapien	647	27	61.4	73	2	Q6S9M5_ORYSA	Q6s9ms_oryza sativ
575	28	63.6	1193	2	Q90819_CHICK	Q90819_gallus gall	648	27	61.4	74	1	Q4RA96_TETNG	Q4ra96_tetradodon n
576	28	63.6	1202	1	JAG2_RAT	P97607_rattus norv	649	27	61.4	74	2	MT_CRAVI	P23038_cirsosostrea
577	28	63.6	1213	1	JAG1B_BRARE	Q90Y54_brachydanio	650	27	61.4	74	2	Q8WQ15_OSTED	Q8wq15_ostrea edul
578	28	63.6	1214	2	Q9MMG2_BLAGA	Q9mmg2_blatteila g	651	27	61.4	74	2	Q6ZFLO_BURMA	Q6zfl0_burkholderi
579	28	63.6	1214	2	Q90YD2_XENILA	Q90yd2_xenopus lae	652	27	61.4	75	2	Q6VQPB_CRAVI	Q6vqp8_cirsosostrea
580	28	63.6	1216	2	Q5TEK7_BRARE	Q5tek7_brachydanio	653	27	61.4	75	2	Q9U1N3_CRAVI	Q9u1n3_cirsosostrea
581	28	63.6	1216	2	Q90Y55_BRARE	Q90y55_brachydanio	654	27	61.4	75	2	Q53ZD1_CRAVI	Q53zdl_cirsosostrea
582	28	63.6	1218	1	JAG1_HUMAN	P78504_homo sapien	655	27	61.4	76	2	Q8MUZ7_CRAVI	Q8muz7_cirsosostrea
583	28	63.6	1218	1	JAG1_MOUSE	Q9qxx0_mus musculu	656	27	61.4	76	2	Q6S9RS_ORYSA	Q6s9rs_oryza sativ
584	28	63.6	1218	1	Q4KMR2_HUMAN	Q4kmr2_homo sapien	657	27	61.4	76	2	Q7ULQ4_RHOBA	Q7ulq4_rhodopirell
585	28	63.6	1219	1	JAG1_RAT	Q6j722_rattus norv	658	27	61.4	80	2	Q9BIV4_9BIVA	Q9biv4_cirsosostrea
586	28	63.6	1238	1	JAG1A_HUMAN	Q90Y57_brachydanio	659	27	61.4	83	2	Q7JVC4_DROME	Q7jvc4_drosophila
587	28	63.6	1242	1	JAG1A_BRARE	Q90Y57_brachydanio	660	27	61.4	86	2	Q4GYK6_9TRYX	Q4gyk6_9tryx
588	28	63.6	1245	2	Q9Y7V5_TRITHA	Q9y7v5_trichoderma	661	27	61.4	86	2	Q8CUD4_YERPE	Q8cud4_yersinia pe
589	28	63.6	1246	2	Q6LA40_HUMAN	Q6la40_homo sapien	662	27	61.4	86	2	Q6E889_YERPE	Q6e889_yersinia ps
590	28	63.6	1247	1	JAG2_MOUSE	Q9qy65_mus musculu	663	27	61.4	87	2	Q54212_STRDR	Q54212_streptomyce
591	28	63.6	1254	2	Q5TEK8_BRARE	Q5tek8_brachydanio	664	27	61.4	88	2	Q529S6_ORYSA	Q529s6_oryza sativ
592	28	63.6	1254	2	Q90Y56_BRARE	Q90y56_brachydanio	665	27	61.4	89	2	Q94376_CABEL	Q94376_caenorhabdi
593	28	63.6	1254	2	Q9YHU2_BRARE	Q9yhu2_brachydanio	666	27	61.4	90	2	Q6LA475_ORYSA	Q6la475_oryza sativ
594	28	63.6	1318	2	Q5DIC9_HUMAN	Q5dic9_homo sapien	667	27	61.4	91	2	Q7Y227_9CAUD	Q7y227_siki conver
595	28	63.6	1319	2	Q4S599_CABEL	Q4s599_caenorhabdi	668	27	61.4	91	2	Q8S6B4_9CAUD	Q8s6b4_mycobacteri
596	28	63.6	1374	2	Q6LA41_HUMAN	Q6la41_homo sapien	669	27	61.4	92	2	Q4SB76_TETNG	Q4sb76_tetradodon n
597	28	63.6	1403	2	Q9QXG1_MOUSE	Q9qxx1_mus musculu	670	27	61.4	93	2	Q4SCB2_TETNG	Q4scb2_tetradodon n
598	28	63.6	1434	2	Q4I7C1_GIBZE	Q4i7c1_gibberella	671	27	61.4	94	2	Q6VQPS_CRAVI	Q6vqp5_cirsosostrea
599	28	63.6	1446	2	Q5DID0_HUMAN	Q5did0_homo sapien	672	27	61.4	95	2	Q5WQGO_9Z2Z2	Q5wqgo_uncultured
600	28	63.6	1469	2	Q6IGM7_CABER	Q6igm7_caenorhabdi	673	27	61.4	97	2	Q6S6S1_9VIRU	Q6s6s1_lumbo virus
601	28	63.6	1477	2	Q4H3A4_CIOIN	Q4h3a4_ciona intes	674	27	61.4	97	2	Q88472_9VIRU	Q88472_tatyna viru
602	28	63.6	1592	2	Q4IDJ7_GIBZE	Q4idj7_gibberella	675	27	61.4	99	2	Q98751_9VIRU	Q98751_tatyna viru
603	28	63.6	1664	2	Q9TVQ2_CABEL	Q9tvq2_caenorhabdi	676	27	61.4	99	2	Q7UL13_MYCBO	Q7ul13_mycobacteri
604	28	63.6	1691	2	Q4HUY8_GIBZE	Q4huy8_gibberella	677	27	61.4	99	2	Q7D937_MYCTH	Q7d937_mycobacteri
605	28	63.6	1709	2	Q80XK8_MOUSE	Q80xk8_mus musculu	678	27	61.4	100	2	Q7JAV1_ACRTH	Q7jav1_agrobacteri
606	28	63.6	1717	2	Q2S566_SCIMA	Q2s566_schistosoma	679	27	61.4	100	2	Q8R4C3_CICDR	Q8r4c3_cricetulus
607	28	63.6	1805	2	Q63661_RAT	Q63661_rattus norv	680	27	61.4	101	2	Q8BQER_MOUSE	Q8bqer_mus musculu
608	28	63.6	1866	2	Q6GQV6_MOUSE	Q6gqv6_mus musculu	681	27	61.4	103	2	Q9AUB8_9LILI	Q9aub8_copeinicia
609	28	63.6	1896	2	Q84MR5_ORYSA	Q84mr5_oryza sativ	682	27	61.4	104	2	Q9NCR5_9CUCU	Q9ncr5_dendroides
610	28	63.6	1968	2	Q8XOC5_NEUCR	Q8xoc5_neuropora	683	27	61.4	105	2	Q6B715_RABIT	Q6b715_oryctolagus
611	28	63.6	2113	2	Q7XW87_ORYSA	Q7xw87_oryza sativ	684	27	61.4	105	2	Q29198_PIG	Q29198_sus scrofa
612	28	63.6	2117	2	Q9H4D8_HUMAN	Q9h4d8_homo sapien	685	27	61.4	107	2	Q9NG19_CRACI	Q9ng19_cirsosostrea
613	28	63.6	2125	2	Q7QWMI_GIALA	Q7qwmi_giardia lam	686	27	61.4	110	2	Q6IWR8_9CALI	Q6iwr8_porcine ent
614	28	63.6	2167	2	Q9H481_HUMAN	Q9h481_homo sapien	687	27	61.4	110	2	Q6IWR9_9CALI	Q6iwr9_porcine ent
615	28	63.6	2169	2	Q9NY09_HUMAN	Q9ny09_homo sapien	688	27	61.4	112	2	O16120_TENMO	O16120_tenebrio mo

689	27	61.4	112	2	Q9U745_TENNO	Q9U745_tenebrio mo	762	27	61.4	135	2	Q8S4D5_9LILI	Q8S4D5_physokentia
690	27	61.4	112	2	Q9U746_TENNO	Q9U746_tenebrio mo	763	27	61.4	135	2	Q8S4B8_HYOLA	Q8S4B8_hyophorbe 1
691	27	61.4	112	2	Q9U747_TENNO	Q9U747_tenebrio mo	764	27	61.4	135	2	Q9AU72_9LILI	Q9AU72_pylchospem
692	27	61.4	112	2	Q9U748_TENNO	Q9U748_tenebrio mo	765	27	61.4	135	2	Q6AB87_PROAC	Q6AB87_propionibac
693	27	61.4	113	2	Q5SNV5_ORYSA	Q5snas oryza sativ	766	27	61.4	136	2	Q7X6D0_ORYSA	Q7X6D0_oryza sativ
694	27	61.4	115	1	PLAC8_HUMAN	Q9afit1 homo sapien	767	27	61.4	136	2	Q9AU68_9LILI	Q9AU68_cyphosperma
695	27	61.4	115	2	Q9REK4_HUMAN	Q96ej4 homo sapien	768	27	61.4	136	2	Q6B941_9PAPI	Q6B941_bovine papil
696	27	61.4	115	2	Q9REK4_PONPY	Q5rek4 pongo pygma	769	27	61.4	137	2	Q8BND5_HUMAN	Q8bnds homo sapien
697	27	61.4	118	2	Q9S979_CRAAB	Q9s979 crambe abys	770	27	61.4	137	2	Q9AU67_9LILI	Q9AU67_rectiphiala
698	27	61.4	118	2	Q68531_PSRAC	Q68531 pseudomonas	771	27	61.4	137	2	Q9AU68_9LILI	Q9AU68_sacchara ex
699	27	61.4	118	2	Q68020_PSEAE	Q68020 pseudomonas	772	27	61.4	137	2	Q9AU69_9LILI	Q9AU69_rhopalosytl
700	27	61.4	119	2	Q7UMM3_RHOBA	Q7umm3 rhodopirell	773	27	61.4	137	2	Q9AU70_9LILI	Q9AU70_reinhardtia
701	27	61.4	119	2	Q8BJR5_MOUSE	Q8bjr5 mus musculi	774	27	61.4	137	2	Q9AU71_RAVRI	Q9AU71_ravenea tiv
702	27	61.4	120	1	VE4_HP42	P27225 human papil	775	27	61.4	137	2	Q9AU73_PSESC	Q9AU73_pseudophen
703	27	61.4	120	2	Q9YBR9_AERPE	Q9ybr9 aeropyrum p	776	27	61.4	137	2	Q9AU75_9LILI	Q9AU75_plafeteta e
704	27	61.4	120	2	Q5GCM8_GCIUCU	Q5gcm8 microdera d	777	27	61.4	137	2	Q9AU77_9LILI	Q9AU77_linospadix
705	27	61.4	120	2	Q5G6W9_9CUCU	Q5g6w9 microdera d	778	27	61.4	137	2	Q9AU78_9LILI	Q9AU78_ignanura wa
706	27	61.4	120	2	Q8L930_ARATH	Q8l930 arabidopsis	779	27	61.4	137	2	Q9AU82_DYPLE	Q9AU82_dypsis lept
707	27	61.4	121	1	RNPA_NEIG1	Q5fw43 neisseria g	780	27	61.4	137	2	Q9AU83_9LILI	Q9AU83_dicyosperm
708	27	61.4	121	1	RNPA_NEIG1	Q9jw46 neisseria m	781	27	61.4	137	2	Q9AU84_9LILI	Q9AU84_deckenia no
709	27	61.4	121	1	RNPA_NEIMB	Q9jw46 neisseria m	782	27	61.4	137	2	Q9AU85_9LILI	Q9AU85_cytosactachy
710	27	61.4	121	2	Q7PGT7_ANGGA	Q7pgt7 anopheles g	783	27	61.4	137	2	Q9AU89_CHAHU	Q9AU89_chamaeops
711	27	61.4	121	2	Q5X5B3_LEGPA	Q5x5b3 legionella	784	27	61.4	137	2	Q9AU92_CARMT	Q9AU92_caryota mit
712	27	61.4	123	2	Q7QF28_ANGGA	Q7qf28 anopheles g	785	27	61.4	137	2	Q9AU93_9LILI	Q9AU93_brongnariari
713	27	61.4	123	2	Q45261_CABEL	Q45261 caenorhabd	786	27	61.4	137	2	Q9AU94_9LILI	Q9AU94_bentitockia
714	27	61.4	124	2	Q16121_TENNO	Q16121 tenebrio mo	787	27	61.4	137	2	Q9AU95_9LILI	Q9AU95_beccatiopho
715	27	61.4	124	2	Q9U744_TENNO	Q9U744 tenebrio mo	788	27	61.4	137	2	Q69GY1_9PAPI	Q69gy1 bovine papil
716	27	61.4	125	2	Q6DLX5_TENNO	Q6dlx5 tenebrio mo	789	27	61.4	138	2	Q9AU81_9LILI	Q9AU81_dypsis lute
717	27	61.4	125	2	Q8G977_MOUSE	Q8g977 mus musculi	790	27	61.4	143	2	Q5KOP2_ORYSA	Q5kpf2 oryza sativ
718	27	61.4	125	2	Q9AU74_9LILI	Q9au74 prestoea ca	791	27	61.4	143	2	Q8QNU6_9PAPI	Q8qnu6 human papil
719	27	61.4	127	2	Q9AU76_9LILI	Q9au76 phoenicopho	792	27	61.4	144	2	Q8KLU7_9222Z	Q8klj7 uncultured
720	27	61.4	128	2	Q9YDS2_AERPE	Q9yds2 aeropyrum p	793	27	61.4	144	2	Q6S4F5_9PAPI	Q6s4f5 human papil
721	27	61.4	128	2	Q8BNU7_MOUSE	Q8bnu7 mus musculi	794	27	61.4	144	2	Q6S413_9PAPI	Q6s413 human papil
722	27	61.4	129	2	Q9AU90_CHACE	Q9au90 chamaedorea	795	27	61.4	144	2	Q80K25_9PAPI	Q80k25 human papil
723	27	61.4	130	2	Q8S4D3_9LILI	Q8s4d3 rhopalobias	796	27	61.4	144	2	Q8J5K9_9PAPI	Q8j5k9 chimpanzee
724	27	61.4	130	2	Q8S4D9_9LILI	Q8s4d9 normandya n	797	27	61.4	144	2	Q8QNU7_9PAPI	Q8qnu7 human papil
725	27	61.4	130	2	Q9T040_ARATH	Q9t040 arabidopsis	798	27	61.4	144	2	Q8QR28_9PAPI	Q8qr28 human papil
726	27	61.4	130	2	Q9BGV7_RHILIO	Q9bgv7 rhizobium l	799	27	61.4	144	2	Q8QR29_9PAPI	Q8qr29 human papil
727	27	61.4	131	2	Q5T7M1_HUMAN	Q5t7m1 homo sapien	800	27	61.4	144	2	Q8VLR6_9PAPI	Q8vlr6 human papil
728	27	61.4	131	2	Q7XJY2_ORYSA	Q7xjy2 oryza sativ	801	27	61.4	144	2	Q9JGX9_9PAPI	Q9jgx9 human papil
729	27	61.4	131	2	Q9AU79_9LILI	Q9au79 heterospath	802	27	61.4	145	1	PA21B_LARSE	PA0611 laticauda s
730	27	61.4	132	2	Q8S4F5_9LILI	Q8s4f5 actinorhytl	803	27	61.4	145	2	Q6VOP2_CRAVI	Q6vop2 crassostrea
731	27	61.4	132	2	Q9AU97_9LILI	Q9au97 bacillus gas	804	27	61.4	145	2	Q7D250_AGRTS	Q7d250 agrobacteri
732	27	61.4	133	2	Q6ZBR3_ORYSA	Q6zbr3 oryza sativ	805	27	61.4	145	2	Q8UJZ6_9PAPI	Q8juz6 macaca faec
733	27	61.4	133	2	Q9AU80_9LILI	Q9au80 sclerosperm	806	27	61.4	145	2	Q8V919_9PAPI	Q8v919 human papil
734	27	61.4	133	2	Q9AU80_9LILI	Q9au80 tetradon n	807	27	61.4	145	2	Q6ZOW3_HUMAN	Q6zow3 homo sapien
735	27	61.4	134	2	Q4SDH4_TENNG	Q4sdh4 tetradon n	808	27	61.4	146	2	Q80K31_9PAPI	Q80k31 human papil
736	27	61.4	134	2	Q5Q0B1_ARATH	Q5q0b1 arabidopsis	809	27	61.4	146	2	Q8V9J1_9PAPI	Q8v9j1 human papil
737	27	61.4	134	2	Q8S4C6_9LILI	Q8s4c6 wellia regi	810	27	61.4	146	2	Q9WGF3_9PAPI	Q9wgf3 human papil
738	27	61.4	134	2	Q8S4C8_9LILI	Q8s4c8 vereschafel	811	27	61.4	147	2	Q8JSL6_9PAPI	Q8jsl6 chimpanzee
739	27	61.4	134	2	Q8S4C9_9LILI	Q8s4c9 sommieria e	812	27	61.4	147	2	Q8UJZ3_9PAPI	Q8juz3 macaca faec
740	27	61.4	134	2	Q8S4D0_9LILI	Q8s4d0 sclerosperm	813	27	61.4	147	2	Q9DMV6_9PAPI	Q9dmv6 human papil
741	27	61.4	134	2	Q8S4D1_ROTRE	Q8s4d1 satkentia	814	27	61.4	148	2	Q9DMC8_9CUCU	Q9dmc8 dendroides
742	27	61.4	134	2	Q8S4D2_9LILI	Q8s4d2 rosetonea r	815	27	61.4	148	2	O16122_TENNO	O16122_tenebrio mo
743	27	61.4	134	2	Q8S4D4_9LILI	Q8s4d4 roschertia m	816	27	61.4	148	2	Q6H5T8_ORYSA	Q6h5t8 oryza sativ
744	27	61.4	134	2	Q8S4D6_9LILI	Q8s4d6 podococcus	817	27	61.4	148	2	Q9ERK8_MOUSE	Q9erk8 mus musculi
745	27	61.4	134	2	Q8S4D7_9LILI	Q8s4d7 pelagodoxa	818	27	61.4	149	2	Q6VOP3_CRAVI	Q6vop3 crassostrea
746	27	61.4	134	2	Q8S4D8_9LILI	Q8s4d8 oncosperma	819	27	61.4	149	2	Q6VOP4_CRAVI	Q6vop4 crassostrea
747	27	61.4	134	2	Q8S4E1_9LILI	Q8s4e1 nenga pumil	820	27	61.4	150	2	Q7RU43_NEUCR	Q7ru43 neurospora
748	27	61.4	134	2	Q8S4E2_9LILI	Q8s4e2 masocala mad	821	27	61.4	151	1	NGB_PIG	NGB_pig sus scrofa
749	27	61.4	134	2	Q8S4E3_9LILI	Q8s4e3 macrojaya d	822	27	61.4	152	2	Q4PIW5_CABEL	Q4piw5 caenorhabd
750	27	61.4	134	2	Q8S4E5_9LILI	Q8s4e5 lodolicea ma	823	27	61.4	154	2	Q7R3E7_GIALA	Q7r3e7 gibberella
751	27	61.4	134	2	Q8S4E6_9LILI	Q8s4e6 leopoldinia	824	27	61.4	155	2	Q4J3J8_GIBBEZ	Q4j3j8 gibberella
752	27	61.4	134	2	Q8S4E7_9LILI	Q8s4e7 lemmrophoen	825	27	61.4	155	2	Q6ZS77_HUMAN	Q6zs77 homo sapien
753	27	61.4	134	2	Q8S4E9_9LILI	Q8s4e9 gytrophyllu	826	27	61.4	155	2	Q5DCT3_SCHJA	Q5dct3 schistosoma
754	27	61.4	134	2	Q8S4F0_9LILI	Q8s4f0 gausbia may	827	27	61.4	155	2	Q4KXK4_PSEFS	Q4kxk4 pseudomonas
755	27	61.4	134	2	Q8S4F2_9LILI	Q8s4f2 cypsis hete	828	27	61.4	155	2	Q5PFYB_SALPA	Q5pfyb salmoneila
756	27	61.4	134	2	Q8S4F3_9LILI	Q8s4f3 carpyosachy	829	27	61.4	155	2	Q8Z916_SALTI	Q8z916 salmoneila
757	27	61.4	134	2	Q8S4F4_9LILI	Q8s4f4 areca cathec	830	27	61.4	157	2	Q4TEFL_BOYING	Q4tefl tetradon n
758	27	61.4	134	2	Q9AU66_BALSE	Q9au66 balaka seem	831	27	61.4	161	2	Q6NTG6_HUMAN	Q6ntg6 homo sapien
759	27	61.4	134	2	Q8UJ11_AGRTS	Q8uj11 agrobacteri	832	27	61.4	163	2	Q6K6S2_ORYSA	Q6k6s2 oryza sativ
760	27	61.4	135	1	H32_XENIA	P02302 xenopus lae	833	27	61.4	164	2	Q4H1M0_9BACT	Q4h1m0 uncultured
761	27	61.4	135	1	H32_XENIA		834	27	61.4	164	2	Q4H1M0_9BACT	

835	27	61.4	167	2	Q5D5H7_9RICK	Q5d5h7 wolbachia e	908	27	61.4	244	1	RS6_BRAFL	O01727 branchiosto
836	27	61.4	167	2	Q8UJL3_9POXV	Q8uyl3 vaccinia vl	909	27	61.4	245	2	O14644_HUMAN	O14644 homo sapien
837	27	61.4	167	2	Q8UJL2_COMXP	Q8ujl2 comoxp viru	910	27	61.4	245	2	O75LR5_ORYSA	O75LR5 oryza sativ
838	27	61.4	167	2	Q72762_COMXP	Q72762 comoxp viru	911	27	61.4	246	1	Y208_METUA	O57661 methanococc
839	27	61.4	168	2	Q60BZ3_ORYSA	Q60BZ3 oryza sativ	912	27	61.4	247	1	RS6_APLCA	O9bmx5 aplysia cal
840	27	61.4	168	2	Q5RQ72_GLUOX	Q5rQ72 gluconobact	913	27	61.4	247	2	Q6SPR3_SOTBN	Q6SPR3 glycine max
841	27	61.4	169	2	Q72768_HUMAN	Q72768 homo sapien	914	27	61.4	247	2	Q4VBAB_RAT	Q4VBAB rattus norv
842	27	61.4	170	2	Q6LGV4_PHOPR	Q6lGV4 phorbacter	915	27	61.4	247	2	Q9DID3_MOUSE	Q9DID3 mus musculu
843	27	61.4	172	1	Y3408_IACPPL	Q88k6 lactobacill	916	27	61.4	247	2	Q91VTI_MOUSE	Q91VTI mus musculu
844	27	61.4	172	1	NUCM_PELSTU	Q79679 pelomedusa	917	27	61.4	247	2	Q9PTD6_CHICK	Q9PTD6 gallus gall
845	27	61.4	173	1	NUCM_PELSTU	Q79679 pelomedusa	918	27	61.4	248	1	RS6_DROME	P29327 drosophila
846	27	61.4	175	2	Q9FM63_ARATH	Q9fm63 arabidopsis	919	27	61.4	248	1	Q4RG40_TETNG	Q4RG40 tetradodon n
847	27	61.4	175	2	Q53P21_ORYSA	Q53P21 oryza sativ	920	27	61.4	249	1	RS6B_ARATH	P51430 arabidopsis
848	27	61.4	176	2	Q4J7H8_SULAC	Q4j7H8 sulfolobus	921	27	61.4	249	1	RS6_CHICK	P47838 gallus gall
849	27	61.4	177	2	Q6I112_DROME	Q6i112 drosophila	922	27	61.4	249	1	RS6_HUMAN	P62753 homo sapien
850	27	61.4	177	2	Q9MSY0_ORYSA	Q9mY0 oryza sativ	923	27	61.4	249	1	RS6_ICTPU	O90Y09 ictalurus p
851	27	61.4	177	2	Q9MSY7_ORYSA	Q9mY7 oryza sativ	924	27	61.4	249	1	RS6_MOUSE	P62754 mus musculu
852	27	61.4	179	2	Q7Y061_ORYSA	Q7Yk61 oryza sativ	925	27	61.4	249	1	RS6_ONCMY	O9Y912 oncorhynch
853	27	61.4	183	1	RM43_MOUSE	Q99A89 mus musculu	926	27	61.4	249	1	RS6_RAT	P62755 rattus norv
854	27	61.4	186	2	Q6AMR6_DROME	Q6amr6 drosophila	927	27	61.4	249	1	RS6_XENLA	P33017 xenopus lae
855	27	61.4	186	2	Q9VAL2_DROME	Q9vAl2 drosophila	928	27	61.4	249	2	O8N6Z7_HUMAN	O8N6Z7 homo sapien
856	27	61.4	188	2	Q9L031_STRCO	Q9l031 streptomyce	929	27	61.4	249	2	O9EDV6_HUMAN	O9EDV6 homo sapien
857	27	61.4	188	2	Q9ERS8_RAT	Q9er8 rattus norv	930	27	61.4	249	2	Q4VB77_HUMAN	O4VB77 homo sapien
858	27	61.4	190	2	Q65637_ARATH	Q65637 arabidopsis	931	27	61.4	249	2	Q5CXZ9_CRYPO	O5cxz9 cryptospori
859	27	61.4	190	2	P70368_MOUSE	P70368 mus musculu	932	27	61.4	249	2	Q5CP03_CRYHO	O5CP03 cryptospori
860	27	61.4	191	2	Q4QFX9_LEIMA	Q4qFX9 leishmania	933	27	61.4	249	2	Q5E995_BOVIN	O5E995 boe taurus
861	27	61.4	192	2	Q54V14_DICDI	Q54V14 dictyostell	934	27	61.4	249	2	Q4R4K6_MACFA	Q4R4K6 macaca fasc
862	27	61.4	193	2	Q7UM46_RHOBA	Q7uW46 rhodospirell	935	27	61.4	249	2	Q7Y0H9_BRANA	Q7Y0H9 brassica na
863	27	61.4	195	2	Q5SBC9_DICDI	Q5Sbc9 dictyostell	936	27	61.4	249	2	Q9JLM3_NOSFU	Q9JLM3 noscic punc
864	27	61.4	196	2	Q8VC17_MOUSE	Q8vc17 mus musculu	937	27	61.4	249	2	O8BT09_MOUSE	O8BT09 mus musculu
865	27	61.4	197	2	Q7M097_VIBRV	Q7M097 vibrio vuln	938	27	61.4	249	2	O5BLK1_MOUSE	O5BLK1 mus musculu
866	27	61.4	199	2	Q61XF7_CABBR	Q61xf7 caenorhabdi	939	27	61.4	249	2	Q7ZYU0_XENLA	Q7ZYU0 xenopus lae
867	27	61.4	200	2	Q6VOP0_CRAVI	Q6vOp0 crassostrea	940	27	61.4	249	2	Q5D086_XENLA	O5D086 xenopus lae
868	27	61.4	201	2	Q8GTL3_TOBAC	O8gTl3 nicotiana t	941	27	61.4	249	2	Q6DHL6_BRAHE	Q6DHL6 brachydano
869	27	61.4	202	2	Q5EHV1_GECUA	Q5eHv1 gecko japon	942	27	61.4	249	2	Q4SDU7_TETNG	Q4SDU7 tetradodon n
870	27	61.4	204	2	Q6VQPI_CRAVI	Q6vQp1 crassostrea	943	27	61.4	249	2	Q6P7K2_XENTR	Q6P7K2 xenopus tro
871	27	61.4	209	2	Q7WMM2_BORBR	Q7wM2 bordetella	944	27	61.4	250	1	RS6A_ARATH	O46484 cor9BT
872	27	61.4	210	2	Q6IGL5_DROME	Q6iGL5 drosophila	945	27	61.4	250	2	O8LH97_ORYSA	O8LH97 oryza sativ
873	27	61.4	210	2	Q4Q3D3_LEIMA	Q4q3D3 leishmania	946	27	61.4	250	2	Q4TH17_TETNG	Q4TH17 tetradodon n
874	27	61.4	210	2	Q6A8A1_LEIIXX	Q6a8A1 leifsonia x	947	27	61.4	251	1	RS6_ASPOP	O9m3V8 aspergilla o
875	27	61.4	211	2	Q8BR56_MOUSE	Q8bR56 mus musculu	948	27	61.4	251	2	O8IRN9_DROME	O8IRN9 drontophila
876	27	61.4	211	2	Q8R037_MOUSE	Q8r037 mus musculu	949	27	61.4	251	2	O04014_MAIJE	O04014 zea mays (m
877	27	61.4	213	2	Q5F3Y3_CHICK	Q5f3Y3 gallus gall	950	27	61.4	251	2	Q9EY50_MAIJE	Q9EY50 zea mays (m
878	27	61.4	214	2	Q7SSU1_NEUCR	Q7sSul neurospora	951	27	61.4	252	2	O5OX16_ENTHI	O5OX16 entamoeba h
879	27	61.4	215	2	Q8VKP8_MYCTU	O8vKf8 mycobacteri	952	27	61.4	253	1	ERMA_CORDI	P14698 corynebacte
880	27	61.4	219	2	Q8P3P8_NEUCR	Q8p3f8 neurospora	953	27	61.4	253	1	RS6_MANSE	O94624 manduca sex
881	27	61.4	220	1	SODP_CAMJE	P53640 campylobact	954	27	61.4	253	1	RS6_SFOPR	O9SV32 epidoptera
882	27	61.4	220	2	Q5HMZ7_CAMJR	Q5hmz7 campylobact	955	27	61.4	253	2	Q5UAN8_BOWMO	Q5UAN8 bombyx mori
883	27	61.4	222	2	Q6A854_PROAC	Q6a854 propionibac	956	27	61.4	253	2	Q46484_COR9T	Q46484 corynebacte
884	27	61.4	223	2	Q4WBL0_ASPPU	Q4wBl0 aspergillus	957	27	61.4	253	2	Q53WT1_CORDI	O53wt1 corynebacte
885	27	61.4	223	2	Q86VYK5_HUMAN	Q86vYk5 homo sapien	958	27	61.4	253	2	Q7BBX6_PROAC	Q7BBX6 propionibac
886	27	61.4	226	2	Q8A4U3_BACTN	Q8A4U3 bacteroides	959	27	61.4	254	2	Q6TAY3_CHICK	Q6TAY3 gallus gall
887	27	61.4	229	2	Q9NKM5_HUMAN	Q9nkm5 homo sapien	960	27	61.4	255	1	TNR9_HUMAN	Q07011 homo sapien
888	27	61.4	229	2	Q57Z89_HUMAN	Q57Z89 homo sapien	961	27	61.4	255	2	Q9H5H2_HUMAN	Q9H5H2 homo sapien
889	27	61.4	231	1	NRC1_YEAST	Q03125 saccharomyc	962	27	61.4	255	2	Q94LS9_ORYSA	Q94LS9 oryza sativ
890	27	61.4	232	2	Q9SKA7_MACFA	Q9sKf7 macaca fasc	963	27	61.4	256	1	TNR9_MOUSE	P20334 mus musculu
891	27	61.4	233	2	Q8BFW6_XANAC	Q8bFw9 xanthomonas	964	27	61.4	256	2	Q7RRG5_PLAFO	Q7RRG5 plasmodium
892	27	61.4	234	2	Q7Z6G5_HUMAN	Q7Z6G5 homo sapien	965	27	61.4	256	2	Q4YU66_PLAAB	Q4YU66 plasmodium
893	27	61.4	234	2	Q6YSD4_ORYSA	Q6ySD4 oryza sativ	966	27	61.4	256	2	Q4XZM0_PLACH	Q4XZM0 plasmodium
894	27	61.4	234	2	Q9X897_STRCO	Q9X897 streptomyce	967	27	61.4	256	2	Q8F0Z4_LEPIN	Q8F0Z4 leptospira
895	27	61.4	235	2	Q90ZF9_CHICK	Q90zf9 gallus gall	968	27	61.4	257	2	Q9AM12_STRAT	Q9Am12 streptomyce
896	27	61.4	236	1	SRD1_HUMAN	Q9uhv0 mus musculu	969	27	61.4	258	2	Q7ZU49_LEPIC	Q7ZU49 leptospira
897	27	61.4	236	1	SRD1_MOUSE	Q9j110 mus musculu	970	27	61.4	258	2	Q4W8J3_RAT	Q4W8J3 rattus norv
898	27	61.4	236	2	Q53GCO_HUMAN	Q53gCO homo sapien	971	27	61.4	258	2	Q4V895_RAT	Q4V895 rattus norv
899	27	61.4	236	2	Q8XWK3_RALUSO	Q8xwk3 talstonia s	972	27	61.4	259	2	Q93519_CABBL	Q93519 caenorhabdi
900	27	61.4	236	2	Q6P771_RAT	Q6p771 rattus norv	973	27	61.4	259	2	Q63PV9_BURPS	Q63PV9 burkholderi
901	27	61.4	237	1	IPDE_DICDI	P22549 dictyostell	974	27	61.4	259	2	Q7NCMS_GLOVI	Q7NCMS gloeobacter
902	27	61.4	237	2	Q54YU3_DICDI	Q54YU3 dictyostell	975	27	61.4	260	2	O8LN62_ORYSA	O8LN62 oryza sativ
903	27	61.4	237	2	Q45804_CABBL	Q45804 caenorhabdi	976	27	61.4	263	2	Q6CXB9_KLUFA	Q6CXB9 kluyveromyc
904	27	61.4	239	1	GAG_AVEU1	P06936 avian endog	977	27	61.4	265	2	Q9A2T7_CAUCR	Q9A2T7 caulobacter
905	27	61.4	239	1	Q4KTD6_SUBDO	Q4Ktd6 suberites d	978	27	61.4	265	2	Q7Z6V5_DBSVA	Q7Z6V5 deuliovoivr
906	27	61.4	239	2	Q4ZF55_CLOPE	Q4zf55 clostridium	979	27	61.4	267	2	O02764_RABIT	O02764 oryctolagus
907	27	61.4	242	2	Q6JWW6_JUROC	Q6jww6 cistopileura	980	27	61.4	270	2	Q75SV8_PELCA	Q75SV8 felis silve

```
981 27 61.4 270 2 Q9ZUN5 ARATH
982 27 61.4 270 2 Q9LH57 ARATH
983 27 61.4 270 2 Q8BX32 MOUSE
984 27 61.4 271 1 TNR4_RAT
985 27 61.4 271 2 Q54RT0 DICI
986 27 61.4 271 2 Q5PQP5 RAT
987 27 61.4 272 1 TNR4_MOUSE
988 27 61.4 272 1 Q6Z5Z9 ORYSA
989 27 61.4 273 2 Q5UCF4 MAIZE
990 27 61.4 274 1 TNR5_CANPA
991 27 61.4 274 1 Q4IVQ6 AZOVI
992 27 61.4 274 2 Q8P3E6_XANCP
993 27 61.4 275 2 Q8BQ83_HUMAN
994 27 61.4 275 2 Q8QW99_MOUSE
995 27 61.4 276 2 Q4IFD8_GIBZE
996 27 61.4 276 2 Q94473 DICI
997 27 61.4 276 2 Q71F55_MOUSE
998 27 61.4 277 1 TNR4_HUMAN
999 27 61.4 277 1 TNR5_HUMAN
1000 27 61.4 277 2 Q5U007_HUMAN
```

## ALIGNMENTS

## RESULT 1

Q98AR9\_RH10 PRELIMINARY; PRT; 192 AA.

```
ID Q98AR9_RH10 PRELIMINARY; PRT; 192 AA.
AC Q98AR9;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE M15880 protein.
GN OrderedLocustNames=mlr5880;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
NCBI_TaxID=381;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti."
RL EMBL: BA000012; BAB52253.1; -; Genomic_DNA.
DR HSSP: Q92FY9; IPTH.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR002347; ADH_short_C2.
DR InterPro: IPR002424; Insect_adh_fam.
DR PRINTS: PR00081; GDHDPH.
DR PRINTS: PR01167; INSADHFAMILY.
DR PRINTS: PR00080; SDRFAMILY.
DR Complete proteome.
KW SEQUENCE 192 AA; 20183 MW; FD2F660D156037BC CRC64;
```

Query Match 81.8%; Score 36; DB 2; Length 192;  
Best Local Similarity 71.4%; Pred. No. 36;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 176 CSPATGC 182

## RESULT 2

```
Q86AK7_DICDI
ID Q86AK7_DICDI PRELIMINARY; PRT; 360 AA.
AC Q86AK7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 10-MAY-2003 (TrEMBLrel. 30, Last sequence update)
DE Similar to Dictyostelium discoidium (Slime mold). Freestalk protein
DE (Hypothetical protein).
ORFNames=DDB0168042; DDB0217347;
GN Dictyostelium discoidium (Slime mold).
OS Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NCBI_TaxID=44689;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafrański K., Pachebat J.A.,
RA Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
RA Abriil J.F., Guigo R., Kump K., Tunggal B., Cox E., Quail M.A.,
RA Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoidium."
RL Nature 418:79-85(2002).
[2]
RN RN NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RC STRAIN=AX4;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
[3]
RN RN NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RC STRAIN=AX4;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
```

```
ID Q8KJ72_RH10 PRELIMINARY; PRT; 546 AA.
AC Q8KJ72;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 10-MAY-2003 (TrEMBLrel. 30, Last sequence update)
DE Similar to Dictyostelium discoidium (Slime mold). Freestalk protein
DE (Hypothetical protein).
ORFNames=DDB0168042; DDB0217347;
GN Dictyostelium discoidium (Slime mold).
OS Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NCBI_TaxID=44689;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafrański K., Pachebat J.A.,
RA Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
RA Abriil J.F., Guigo R., Kump K., Tunggal B., Cox E., Quail M.A.,
RA Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoidium."
RL Nature 418:79-85(2002).
[2]
RN RN NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RC STRAIN=AX4;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
```

Query Match 81.8%; Score 36; DB 2; Length 360;  
Best Local Similarity 71.4%; Pred. No. 64;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 211 CSPSTGC 217

## RESULT 3

Q8KJ72\_RH10 PRELIMINARY; PRT; 546 AA.

AC O8KJ72;  
 DT 01-OCT-2002 (TReMBLrel. 22, Created)  
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE PUTATIVE SHORT-CHAIN TYPE DEHYDROGENASE/REDUCTASE PROTEIN.  
 GN Name=msj323;  
 OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 CC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxId=381;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=RTA;  
 RX MEDLINE=21999272; PubMed=12003951;  
 RX DOI=10.1128/JB.184.11.3086-3095.2002;  
 RA Sullivan J.T., Eltiot R.M., Fleetwood D.J., McCallum N.G., Rosbach U.,  
 RA Brown S.D., Elliott R.M., Fleetwood D.J., McCallum N.G., Rosbach U.,  
 RA Stuart G.S., Weaver J.E., Webb R.J., de Bruijn F.J., Ransom C.W.;  
 RT "Comparative sequence analysis of the symbiosis island of  
 RT Mesorhizobium loti strain RTA."  
 RL J. Bacteriol. 184:3086-3095(2002).  
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases  
 CC (SPR) family.  
 CC EMBL: AL672114; CAD31361.1; -, Genomic\_DNA.  
 DR HSP; P50163; ZAE1.  
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002198; Adh\_short.  
 DR InterPro; IPR002347; Adh\_short\_C2.  
 DR InterPro; IPR002424; Insect\_adh\_fam.  
 DR Pfam; PF00106; adh\_short; 2.  
 DR PRINTS; PR00081; GDHRD.  
 DR PRINTS; PR01167; INSAHPFAMILY.  
 DR PRINTS; PR00080; SDRPFAMILY.  
 DR PROSITE; PS00061; ADH\_SHORT; 1.  
 KW Oxidoreductase.  
 SQ SEQUENCE 546 AA; 56900 MW; 091D2EFDE8E55A9C CRC64;  
 Query Match 81.8%; Score 36; DB 2; Length 546;  
 Best Local Similarity 71.4%; Pred. No. 94;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 CXPXTGC 10  
 DB 530 CSPATGC 536  
 RESULT 4  
 ID 054X44 DICI1 PRELIMINARY; PRT; 597 AA.  
 AC 054X44;  
 DT 13-SEP-2005 (TReMBLrel. 31, Created)  
 DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=DD80205657;  
 OS Dictyostelium discoideum (Slime mold).  
 CC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 OX NCBI\_TaxId=44689;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=AX4;  
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
 RA Sugang R., Berriam M., Song J., Olsen R., Szatranski K., Xu Q.,  
 RA Tunggal B., Kummerfeld S., Magera M., Konfortov B.A., Rivoero F.,  
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,  
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
 RA Kerionou A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,  
 RA Farroher P., Desany B., Just E., Morio T., Rost R., Churcher C.,  
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,  
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,  
 RA Wardrop A., Felder M., Thangaveilu M., Johnson D., Knights A.,

RA Louesged H., Mungall K., Oliver K., Price C., Quail M.A.,  
 RA Urushihara H., Hernandez J., Rabinowitz E., Steffen D., Sanders M.,  
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
 RA Shalasky G., Schreiner M., Weinstock G., Rosenthal A., Cox E.C.,  
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
 RA Williams J., Dear P.H., Noegel A.A., Bärrell B., Kuspa A.;  
 RT "The genome of the social amoeba Dictyostelium discoideum."  
 RL Nature 0:0-0(2005).  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBS whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAF10100059; EAL67758.1; -, Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 597 AA; 61409 MW; AA97FID0234D336C CRC64;  
 Query Match 81.8%; Score 36; DB 2; Length 597;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 CXPXTGC 10  
 DB 500 CSPPTGC 506  
 RESULT 5  
 ID 07M4J3 DICI1 PRELIMINARY; PRT; 601 AA.  
 AC 07M4J3;  
 DT 01-MAR-2004 (TReMBLrel. 26, Created)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE DIF-induced prestalk pbd63 protein precursor (fragments).  
 OS Dictyostelium discoideum (Slime mold).  
 CC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 OX NCBI\_TaxId=44689;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=87187613; PubMed=3568124; DOI=10.1016/0092-8674(87)90559-9;  
 RA Williams J.G., Ceccarelli A., McRobbie S., Mahubani H., Kay R.R.,  
 RA Early A., Berks M., Jermyn K.A.;  
 RT "Direct induction of Dictyostelium prestalk gene expression by DIF  
 RT provides evidence that DIF is a morphogen."  
 RL Cell 49:185-192(1987).  
 DR PIR; A27020; A27020.  
 DR InterPro; IPR001673; S\_mold\_repeat.  
 DR Pfam; PF00526; Dicty\_CTDC; 19.  
 FT NON\_TER 601  
 SQ SEQUENCE 601 AA; 63359 MW; 7D4433616CDAC438 CRC64;  
 Query Match 81.8%; Score 36; DB 2; Length 601;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 CXPXTGC 10  
 DB 245 CSPSTGC 251  
 RESULT 6  
 ID 0627A0 CAEBR PRELIMINARY; PRT; 608 AA.  
 AC 0627A0;  
 DT 25-OCT-2004 (TReMBLrel. 28, Created)  
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)  
 DE Hypothetical protein CBG00767 (Fragment).  
 GN Name=CBG00767;  
 OS Caenorhabditis briggsae.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxId=6238;  
 RN [1]



RP NUCLEOTIDE SEQUENCE.  
 RG The C. briggsae Sequencing Consortium;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL: CAAC01000006; CAES7750.1; -; Genomic\_DNA.  
 DR InterPro: IPR006209; EGF-like.  
 DR InterPro: IPR006150; Worm\_repeat\_1.  
 DR SMART: SM00289; WRI; 6.  
 DR SMART: PS01186; EGF\_2; UNKNOWN\_2.  
 DR Hypothetical protein.  
 KW NON TER  
 FT 608 608  
 SQ SEQUENCE 608 AA; 65898 MW; F09E2470D7BE6393 CRC64;

Query Match 81.8%; Score 36; DB 2; Length 608;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10  
 Db 205 CAPSTGC 211

RESULT 7  
 OSAYG2 DICI1 PRELIMINARY; PRT; 1710 AA.  
 AC 054YG2;  
 DT 13-SEP-2005 (TREMBlrel. 31, Created)  
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)  
 DE Extracellular matrix protein STR430.  
 GN Name=ecma; ORFNames=DBG0220137;  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=AX4;  
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
 RA Sungang R., Bertrian M., Song J., Olsen R., Szatranski K., Xu Q.,  
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,  
 RA Banker A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,  
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
 RA Kertornou A., Nie X., Hall N., Anjard C., Hemphill L., Baeson N.,  
 RA Farbrother P., Desany B., Just E., Morio T., Roet R., Churchhead C.,  
 RA Cooper J., Haydock S., Van Driessche N., Cronin A., Goodhead I.,  
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,  
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,  
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,  
 RA Louisgeed H., Mungall K., Oliver K., Price C., Quail M.A.,  
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,  
 RA Ma J., Kohara Y., Sharp S., Simmonds W., Spiegler S., Tivey A.,  
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,  
 RA Chiplom R.L., Gibbs R., Loomis W.F., Platzner M., Kay R.R.,  
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;  
 RT "The genome of the social amoeba Dictyostelium discoideum."  
 RL Nature 0:0-0 (2005).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL: AAF101000052; EAL6809.1; -; Genomic\_DNA.  
 KW Matrix protein.  
 SQ SEQUENCE 1710 AA; 17828 MW; 008C47B85EFD7D2 CRC64;

Query Match 81.8%; Score 36; DB 2; Length 1710;  
 Best Local Similarity 71.4%; Pred. No. 2.7e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10  
 Db 245 CSPSTGC 251

RESULT 8  
 OS5K22\_MACFA  
 ID 095KD2\_MACFA PRELIMINARY; PRT; 148 AA.  
 AC 095KD2;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 OC Cercopithecoidea; Cercopithecoinae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Medulla oblongata;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB062948; BAB60737.1; -; mRNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 148 AA; 15473 MW; C8D2AA301E0C8191 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 148;  
 Best Local Similarity 71.4%; Pred. No. 43;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10  
 Db 74 CGPSTGC 80

RESULT 9  
 OS9N018\_MACFA  
 ID 09N018\_MACFA PRELIMINARY; PRT; 148 AA.  
 AC 09N018;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 OC Cercopithecoidea; Cercopithecoinae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Cerebellum cortex;  
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
 RA Suzuki Y., Sugano S., Hashimoto K.;  
 RT "Isolation of full-length cDNA clones from macaque brain cDNA  
 RT libraries."  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB046629; BAB03547.1; -; mRNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 148 AA; 15495 MW; 6242BCB430C6E13B CRC64;

Query Match 79.5%; Score 35; DB 2; Length 148;  
 Best Local Similarity 71.4%; Pred. No. 43;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10  
 Db 74 CGPSTGC 80

RESULT 10  
 OS5BT6 DICI1  
 ID 055BT6 DICI1 PRELIMINARY; PRT; 370 AA.  
 AC 055BT6;  
 DT 13-SEP-2005 (TREMBlrel. 31, Created)

DT 13-SEP-2005 (TREMblrel. 31, last sequence update)  
 DT 13-SEP-2005 (TREMblrel. 31, last annotation update)  
 DE Hypothetical protein.  
 OS ORFNames=DD0191005;  
 GN Dictyostellum discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 NC NCB1\_TaxID=44689;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=AX4;  
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
 RA Sugand R., Berriman M., Song J., Olsen R., Szatranski K., Xu Q.,  
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,  
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,  
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
 RA Kerionnou A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,  
 RA Parrotier P., Desany B., Just E., Morio T., Rost R., Churcher C.,  
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,  
 RA Hauser H., James K., Quiles M., Mohan M.B., Salto T., Buchrieser C.,  
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,  
 RA Louised H., Mungall K., Oliver K., Price C., Quail M.A.,  
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,  
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
 RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,  
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,  
 RA "The genome of the social amoeba Dictyostelium discoideum.";  
 RL Nature 0:0-0(2005).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAF10100011; EAL72534.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 370 AA; 40291 MW; FB3DF46B49A751A CRC64;

Query Match 79.5%; Score 35; DB 2; Length 370;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
 Db 293 CNPSTGC 299

RESULT 11  
 Q55BX0 D1CD1 PRELIMINARY; PRT; 370 AA.  
 AC Q55BX0;  
 DT 13-SEP-2005 (TREMblrel. 31, Created)  
 DT 13-SEP-2005 (TREMblrel. 31, last sequence update)  
 DT 13-SEP-2005 (TREMblrel. 31, last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=DD0190973;  
 OS Dictyostellum discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 NC NCB1\_TaxID=44689;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=AX4;  
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
 RA Sugand R., Berriman M., Song J., Olsen R., Szatranski K., Xu Q.,  
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,  
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,  
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
 RA Kerionnou A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,  
 RA Parrotier P., Desany B., Just E., Morio T., Rost R., Churcher C.,  
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,  
 RA Hauser H., James K., Quiles M., Mohan M.B., Salto T., Buchrieser C.,  
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,  
 RA Louised H., Mungall K., Oliver K., Price C., Quail M.A.,

RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,  
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
 RA Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,  
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,  
 RA "The genome of the social amoeba Dictyostelium discoideum.";  
 RL Nature 0:0-0(2005).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAF10100011; EAL72514.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 370 AA; 40278 MW; 1E904B2F1186E25 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 370;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
 Db 293 CNPSTGC 299

RESULT 12  
 O5KKP3 CRYNE PRELIMINARY; PRT; 466 AA.  
 ID O5KKP3 CRYNE PRELIMINARY;  
 AC O5KKP3;  
 DT 10-MAY-2005 (TREMblrel. 30, Created)  
 DT 10-MAY-2005 (TREMblrel. 30, last sequence update)  
 DT 10-MAY-2005 (TREMblrel. 30, last annotation update)  
 DE Expressed protein.  
 GN ORFNames=CR02330;  
 OS Cryptococcus neoformans var. neoformans JEC21.  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
 CC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.  
 NC NCB1\_TaxID=214684;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=JEC21;  
 RA Loftus B., Amedeo P., Roncaglia P., Vamathavan J., Uterback T.,  
 RA Van Aken S., Fraser C.,  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=JEC21;  
 RX PubMed=15653466; DOI=10.1126/science.1103773;  
 RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,  
 RA Vamathavan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,  
 RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,  
 RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,  
 RA Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krywinski M.I.,  
 RA Kwon-Chung K.J., Lengele K.B., Malet M., Maier W.A., Maier R.B.,  
 RA Mathewson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,  
 RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,  
 RA Suh B.B., Tenney A., Uterback T.R., Wicks B.L., Wortman J.R.,  
 RA Wye N.H., Kronstad J.W., Lodge J.K., Heltman J., Davis R.W.,  
 RA Fraser C.M., Hyman R.W.,  
 RT "The genome of the basidiomycetous yeast and human pathogen  
 RT Cryptococcus neoformans.";  
 RL Science 307:1321-1324(2005).  
 DR EMBL; AB017343; AAM42205.1; -; Genomic\_DNA.  
 DR InterPro; IPR001452; SH3.  
 DR Pfam; PF00019; SH3.1; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 KW Complete proteome.  
 SQ SEQUENCE 466 AA; 50328 MW; 8D287B486E77EE27 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 466;  
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

Db | | | |  
408 CDPSTGC 414

RESULT 13  
O9AOP5\_PSERE PRELIMINARY;

PRT; 571 AA.

AC O9AOP5\_PSERE PRELIMINARY; PRT; 571 AA.  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Putative ABC transporter subunit.  
GN Name=ORF31;  
OS Pseudomonas resinovorans.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=53412;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=CA10;  
RX MEDLINE=21264379; PubMed=11371531;  
DOI=10.1128/JB.183.12.3663-3679.2001;  
RA Nojiri H., Sekiguchi H., Maeda K., Urata M., Nakai S., Yoshida T.,  
Habe H., Omori T.;  
RT "Genetic characterization and evolutionary implications of car gene  
cluster in carbazole-degrader, Pseudomonas sp. strain CA10."  
RL J. Bacteriol. 183:3663-3679(2001).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=CA10;  
RX MEDLINE=97386424; PubMed=9244273;  
Sato S., Onchiyama N., Kimura T., Nojiri H., Yamane H., Omori T.;  
RT "Cloning of genes involved in carbazole degradation of Pseudomonas sp.  
strain CA10: nucleotide sequence of genes and characterization of  
meta-cleavage enzymes and hydrolase."  
RL J. Bacteriol. 179:4841-4849(1997).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=CA10;  
RX MEDLINE=97386425; PubMed=9244274;  
Sato S., Nam J., Kasuga K., Nojiri H., Yamane H., Omori T.;  
RT "Identification and characterization of genes encoding carbazole 1,9a-  
dioxygenase in Pseudomonas sp. strain CA10."  
RL J. Bacteriol. 179:4850-4858(1997).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=CA10;  
RX PubMed=15466034; DOI=10.1128/JB.186.20.6815-6823.2004;  
Urata M., Miyakoshi M., Kai S., Maeda K., Habe H., Omori T.,  
Yamane H., Nojiri H.;  
RT "transcriptional regulation of the ant operon, encoding two-component  
antitranslate 1,2-dioxygenase, on the carbazole-degradative plasmid  
pCAR1 of Pseudomonas resinovorans strain CA10."  
RL J. Bacteriol. 186:6815-6823(2004).  
DR EMBL; AB047548; BAB32742.1; -; Genomic\_DNA.  
DR GO; GO:0016020; C-membrane; IEA.  
DR GO; GO:0005215; F-transporter activity; IEA.  
DR GO; GO:0006810; P-transport; IEA.  
DR InterPro; IPR001851; Bac\_inmem\_transp.  
DR Pfam; PF02653; BPD\_transp\_2; 2.  
SQ SEQUENCE 571 AA; 60653 MW; 9A885477078C186 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 571;  
Best Local Similarity 71.4%; Pred. NO. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
Db 565 CDPSTGC 571

RESULT 14  
O4SBN7\_TETNG

ID O4SBN7\_TETNG PRELIMINARY; PRT; 997 AA.

AC O4SBN7\_TETNG PRELIMINARY; PRT; 997 AA.  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Chromosome 15 SCAF14667, whole genome shotgun sequence.  
DE (Fragment).  
GN ORFNames=GSTENG00020889001;  
OS Tetraodon nigroviridis (Green puffer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;  
OC Tetraodontidae; Tetraodontidae; Tetraodon.  
OX NCBI\_TaxID=99883;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
Maucci E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
Anthouard V., Jubin C., Catalicio L., Poulain J., De Bernardis V.,  
Biemont C., Skallil Z., Cottolico L., Coutanceau J.P., Gouzy J.,  
Cruaud C., Duprat S., Brothier P., Coutanceau J.P., Gouzy J.,  
Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,  
Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
Laudet V., Schachler V., Queller F., Saurin W., Scarpelli C.,  
Mincker P., Lander E.S., Weissbach J., Roest Crolius H.;  
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
the early vertebrate proto-karyotype."  
RL Nature 431:946-957(2004).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC Genoscope; Whitehead Institute Centre for Genome Research;  
Submitted (Feb-2004) to the EMBL/GenBank/DBJ databases.  
RL -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; CA601014667; CAG01945.1; -; Genomic\_DNA.  
FT NON\_TER 1 997  
FT NON\_TER 1 997  
SQ SEQUENCE 997 AA; 114462 MW; 60278BFED562DEBC CRC64;

Query Match 79.5%; Score 35; DB 2; Length 997;  
Best Local Similarity 71.4%; Pred. NO. 8.6e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
Db 293 CDPSTGC 299

RESULT 15  
O4SF52\_TETNG

PRT; 1018 AA.

AC O4SF52\_TETNG PRELIMINARY; PRT; 1018 AA.  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Chromosome undetermined SCAF14608, whole genome shotgun sequence.  
DE (Fragment).  
GN ORFNames=GSTENG00019252001;  
OS Tetraodon nigroviridis (Green puffer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;  
OC Tetraodontidae; Tetraodontidae; Tetraodon.  
OX NCBI\_TaxID=99883;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
Maucci E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,

RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Jubin C., Castellino L., Katinka M., Vacherie B.,  
 RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bernardis V.,  
 RA Crnaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,  
 RA Parra G., Lardier G., Chapelle C., McKernan K.J., McSwan P., Bosak S.,  
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
 RA Landt V., Schacher V., Quetier F., Saurin W., Scarpelli C.,  
 RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.,  
 RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
 the early vertebrate proto-karyotype.";  
 RT Nature 431:946-957(2004).  
 RL [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Genoscope, Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL, CA001014608; CAG00730.1; -; Genomic\_DNA.  
 DR InterPro: IPR006209; EGF like.  
 DR InterPro: IPR002049; Laminin EGF.  
 DR InterPro: IPR008211; Laminin\_N.  
 DR InterPro: IPR001134; Netrin\_C.  
 DR Pfam: PF00553; Laminin\_EGF; 3.  
 DR Pfam: PF01759; NTR; 1.  
 DR SMART: SM00643; C345C; 1.  
 DR SMART: SM00180; EGF Lam; 3.  
 DR SMART: SM00136; LAMNT; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_2.  
 DR PROSITE: PS01248; EGF\_LAM\_1; 3.  
 DR PROSITE: PS50027; EGF\_LAM\_2; 3.  
 DR PROSITE: PS5117; LAMININ\_NTR; 1.  
 DR PROSITE: PS50189; NTR; 1.  
 DR Laminin EGF-like domain.  
 KM Laminin EGF-like domain.  
 FT NON\_TER  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 1018 AA; 113875 MW; A47C021431A85FF4 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 1018;  
 Best Local Similarity 71.4%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
 DB 770 CSBITGC 776

RESULT 16  
 ID 045795.TETNG PRELIMINARY; PRT; 1839 AA.  
 AC 045795;  
 DT 13-SEP-2005 (TREMBlrel. 31, Created)  
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)  
 DE Chromosome 1 SCAF14716, whole genome shotgun sequence.  
 DE (Fragment).  
 GN ORFNames=GSTENG00022912001;  
 OS Tetraodon nigroviridis (Green puffer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthopterygii; Acanthopterygii; Percomorphi; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Tetraodon.  
 OC NCBI\_TaxID=99883;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
 RA Mauceli E., Bouneau L., Fischer C., Orouf-Coetz C., Bernot A.,  
 RA Nicard S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Seguens B.,  
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Jubin C., Castellino L., Katinka M., Vacherie B.,  
 RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bernardis V.,  
 RA Crnaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,

RA Parra G., Lardier G., Chapelle C., McKernan K.J., McSwan P., Bosak S.,  
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
 RA Landt V., Schacher V., Quetier F., Saurin W., Scarpelli C.,  
 RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.,  
 RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
 the early vertebrate proto-karyotype.";  
 RT Nature 431:946-957(2004).  
 RL [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Genoscope, Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL, CA001014716; CAG03487.1; -; Genomic\_DNA.  
 DR InterPro: IPR001952; ALK\_phosphatase.  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR003887; LEW.  
 DR Pfam: PF00245; ALK\_phosphatase; 2.  
 DR Pfam: PF03020; LEW; 1.  
 DR PRINTS: PR00113; ALKPHPTASE.  
 DR SMART: SM00098; ALKPC; 2.  
 DR SMART: SM00540; LEW; 1.  
 DR PROSITE: PS00123; ALKALINE\_PHOSPHATASE; UNKNOWN\_1.  
 DR PROSITE: PS50954; LEW; 1.  
 DR Magnesium; Zinc.  
 KM Magnesium; Zinc.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 1839 AA; 203924 MW; 05A14C4AAEBC7BC9 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 1839;  
 Best Local Similarity 71.4%; Pred. No. 4.6e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
 DB 1418 CTPGTGC 1424

RESULT 17  
 ID 072607.HUMAN PRELIMINARY; PRT; 113 AA.  
 AC 072607;  
 DT 01-OCT-2003 (TREMBlrel. 25, Created)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE BIF2B5 protein.  
 DE BIF2B5 protein.  
 GN Name=BIF2B5; GN  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;  
 OC Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Lung;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Pelting R.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheefel C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Dlatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stalderon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshynski S., Carinci P., Prange C.,  
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,

```

RA Schnerch A., Schein J.E., Jones S.J.M., Maria M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RG NIH MGC Project;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC053646; AAH53646.1; -; mRNA.
SQ SEQUENCE 113 AA; 11613 MW; 8624A42297FAA3FF CRC64;

Query Match
Best Local Similarity 77.3%; Score 34; DB 2; Length 113;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 103 CYPKTGC 109

RESULT 18
PA2X NOTSC STANDARD; PRT; 145 AA.
ID PA2X NOTSC
AC P20146;
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable phospholipase A2 precursor (BC 3.1.1.4) (Phosphatidylcholine
2-acylhydrolase).
OS Nectechis scutatus scutatus (Mainland tiger snake) (Common tiger
snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodermata; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Acanthophinae; Nectechis.
OX NCBI_TaxId=70142;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=venom gland;
RA Duncanson F.;
RL Submitted (JAN-1989) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-
acyl groups in 3-sn-phosphoglycerides.
CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-
acylglycerophosphocholine + a carboxylate. (By similarity).
CC -1- COFACTOR: Binds 1 calcium ion per subunit. (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the phospholipase A2 family. Group I
subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X14043; CAA32201.1; -; mRNA.
DR PIR; S07983; S07983.
DR HSSP; P00608; 1AE7.
DR InterPro; IPR001211; PhospholipaseA2.
DR PANTHER; PTHR11716; PhospholipaseA2; 1.
DR Pfam; PF00068; PhospholipA2_1; 1.
DR PRINTS; PR00389; PPHPLIPSEA2.
DR ProDom; PD000303; PhospholipaseA2; 1.
DR SMART; SM00085; PA2C; 1.
DR PROSITE; PS00119; PA2_ASP; 1.
DR PROSITE; PS00118; PA2_HIS; 1.
KW Calcium: Hydrolase; Lipid degradation; Metal-binding;
KW Multigene family; Signal.
FT SIGNAL 1 21 Potential.
FT PROPEP 22 27 Potential.
FT CHAIN 28 145 Probable phospholipase A2.
FT ACT_SITE 75 75 By similarity.

```

```

FT ACT_SITE 119 119 By similarity.
FT METAL 55 55 Calcium (via carbonyl oxygen) (By
FT METAL 57 57 Calcium (via carbonyl oxygen) (By
FT METAL 57 57 similarity).
FT METAL 76 76 Calcium (By similarity).
FT DISULFID 38 98 By similarity.
FT DISULFID 54 144 By similarity.
FT DISULFID 56 72 By similarity.
FT DISULFID 71 125 By similarity.
FT DISULFID 78 118 By similarity.
FT DISULFID 87 111 By similarity.
FT DISULFID 105 116 By similarity.
SQ SEQUENCE 145 AA; 16002 MW; 38E36029AE5FAC9 CRC64;

Query Match
Best Local Similarity 77.3%; Score 34; DB 1; Length 145;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 105 CDPETGC 111

RESULT 19
Q7U7V9 SYNXP PRELIMINARY; PRT; 167 AA.
ID Q7U7V9 SYNXP
AC Q7U7V9;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical precursor.
CN OrderedCusNames=SYNM0871;
OS Synecococcus sp. (Strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxId=84588;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahamsa B., Larimer F.W., Land M.L., Hauser L.,
RA Chaitin P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
RA Paulsen I.T., Dutresne A., Patensky F., Webb E.A., Waterbury J.;
RL "The genome of a motile marine Synecococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569691; CAE07386.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein; Signal.
FT SIGNAL 1 22 Potential.
SQ SEQUENCE 167 AA; 18419 MW; 0B3AFB830BCF971C CRC64;

Query Match
Best Local Similarity 77.3%; Score 34; DB 2; Length 167;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
DB 133 CYPGTGC 139

RESULT 20
Q54YF0 D1CDI PRELIMINARY; PRT; 296 AA.
ID Q54YF0 D1CDI
AC Q54YF0;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DDB0218080;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxId=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=AX4;

```

RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
 RA Sucgang R., Berriman M., Song J., Olsen R., Szatranski K., Xu Q.,  
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,  
 RA Banker A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,  
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,  
 RA Fairbrother P., Deany B., Just E., Morio T., Rost R., Churcher C.,  
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,  
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,  
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,  
 RA Louised H., Mungall K., Oliver K., Price C., Quail M.A., Sanders M.,  
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D.,  
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,  
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,  
 RA "The genome of the social amoeba Dictyostelium discoideum.";  
 RT Nature 0:0-0(2005).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL: AA0100052; EMBL6495.1; -; Genomic\_DNA.  
 DR Hypothetical protein.  
 KW KW  
 SQ SEQUENCE 296 AA; 33010 MW; A741D9A5252B5479 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 296;  
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
 Db 26 CDPNTGC 32

RESULT 21  
 O95UV3 ARATH PRELIMINARY; PRT; 312 AA.  
 ID O95UV3 ARATH PRELIMINARY;  
 AC O95UV3;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Hypothetical protein F884.80 (Hypothetical protein AT4g32380).  
 GN Name=F884.80; Synonyms=AT4g32380;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Bevan M., Terry N., Ardiles W., Buysaert C., Daseville R.,  
 RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,  
 RA Villalón R., Gleiden J., Van Montagu M., Hohelsel J., Mewes H.W.,  
 RA Mayer K.F.X., Lemke K., Schueller C.;  
 RA Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Terry N., Ardiles W., Buysaert C., Daseville R., De Clerck R.,  
 RA De Keyser A., Neyt P., Rouze P., Van Den Daele H., Villalón R.,  
 RA Gleiden J., Van Montagu M., Mewes H.W., Lemke K., Mayer K.F.X.,  
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL01581; CAB79955.1; -; Genomic\_DNA.

DR PIR: T05348; T05348.  
 DR GO: GO:004650; F:polysaccharonase activity; IEA.  
 DR GO: GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro: IPR000743; Glyco\_hydro\_28.  
 DR InterPro: IPR006626; PBH.  
 DR Pfam: PF00295; Glyco\_hydro\_28; 2.  
 DR SMART: SM00710; PBH1.3.  
 KW Hypothetical protein.  
 SQ SEQUENCE 312 AA; 34095 MW; E2E70A2622F30BE0 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 312;  
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
 Db 208 CDPNTGC 214

RESULT 22  
 ID 054N14 DICDI PRELIMINARY; PRT; 380 AA.  
 AC 054N14;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=DD80185573;  
 OS Dictyostelium discoideum (slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=AX4;

RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
 RA Sucgang R., Berriman M., Song J., Olsen R., Szatranski K., Xu Q.,  
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,  
 RA Banker A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,  
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,  
 RA Fairbrother P., Deany B., Just E., Morio T., Rost R., Churcher C.,  
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,  
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,  
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,  
 RA Louised H., Mungall K., Oliver K., Price C., Quail M.A., Sanders M.,  
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D.,  
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,  
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,  
 RA "The genome of the social amoeba Dictyostelium discoideum.";  
 RT Nature 0:0-0(2005).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL: AA01000124; EMBL64620.1; -; Genomic\_DNA.  
 DR Hypothetical protein.  
 KW KW  
 SQ SEQUENCE 380 AA; 40195 MW; A40E4A8D346C8BD CRC64;

Query Match 77.3%; Score 34; DB 2; Length 380;  
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
 Db 124 CDPNTGC 130

RESULT 23  
 O6TVH9\_9POXV PRELIMINARY; PRT; 465 AA.  
 ID O6TVH9\_9POXV PRELIMINARY;

```

AC 06TVH9;
DT 05-VUL-2004 (TReMBLrel. 27, Created)
DT 05-VUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-VUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypochemical protein.
OS Bovine papular stomatitis virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Parapoxvirus.
OC NCBI_TaxID=129727;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BV-AR02;
RX PubMed=14671098; DOI=10.1128/JVI.78.1.168-177.2004;
RA Delton G., Tulman E.R., Afonso C.L., Lu Z., Piccone M.E., Kutish G.F.,
RA de la Concha-Bernierillo A., Leimkuhl H.D., Piccone M.E., Kutish G.F.,
RA Rock D.L.;
RT "Genomes of the Parapoxviruses Orf Virus and Bovine Papular
RT Stomatitis Virus."
RL J. Virol. 78:168-177(2004).
DR EMBL: AY386265; AAR98366.1; -; Genomic_DNA.
DR InterPro: IPR007027; Pox_F11.
DR Pfam: PF04943; Pox_F11; 1.
KW Hypochemical protein.
SQ SEQUENCE 465 AA; 51404 MW; BD735469CBED1B79 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 465;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
Db 26 CVPRTGC 32

RESULT 24
O4MMC1 ASPFU PRELIMINARY; PRT; 487 AA.
ID O4MMC1 ASPFU PRELIMINARY;
AC O4MMC1;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Chitin deacetylase, putative.
GN ORFNames=Afu6910430;
OS Aspergillus fumigatus Af393.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OC NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293.
RA Nieman W., Pain A., Anderson M.J., Wortman J., Kim H., Stanley,
RA Airova J., Berriman M., Abe K., Archer D.B., Bernerjo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblum T.V., Fischer R.,
RA Foster N., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Goni K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horluchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulikarni R.,
RA Kumagai T., Latton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Kumajori W.H., May G.S., Miller B.L., Mahmoud Y., Molina M., Monod M.,
RA Mounya I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,
RA Penativa M.A., Petrea M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitz E., Rawlins N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Rouning C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekai F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrell B., Denning D.W.,
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus
RT Aspergillus fumigatus."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an

```

```

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAHP01000006; EAL88893.1; -; Genomic_DNA.
DR InterPro: IPR001002; Chitin_bd_1.
DR InterPro: IPR002509; Polysac_deact.
DR Pfam: PF00187; Chitin_bind_1; 1.
DR Pfam: PF01522; Polysac_deac_1; 1.
DR Pfam: PF000609; Chitin_binding_1; 2.
DR SMART: SM00270; Chitin_bd_1; 2.
DR PROSITE: PS00941; CHIT_BIND_I_2; 2.
KW Chitin-binding.
SQ SEQUENCE 487 AA; 52742 MW; 1386016F0E98F50 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 487;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
Db 241 CLPRTGC 247

RESULT 25
O5ZAZ8 ORYSA PRELIMINARY; PRT; 504 AA.
ID O5ZAZ8 ORYSA PRELIMINARY;
AC O5ZAZ8;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Putative aspartic proteinase nepenthesin II.
GN Name=PO413G02.28;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22337376; PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hishida S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikono M., Ito S., Ito T., Ito Y., Ito Y., Iwabuchi A., Kamiya K.,
RA Kanaeawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Kanchita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takasaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yokawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.-I., Eun M.-Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1."
RL Nature 420:312-316(2002).
DR EMBL: AP003344; BAD53242.1; -; Genomic_DNA.
DR GO: GO:0004194; P:pepsin A activity; IEA.
DR GO: GO:0005508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR001461; Peptidase_A1.
DR Pfam: PF00026; Asp_1.
DR PRINTS: PR00792; PEPsin.
SQ SEQUENCE 504 AA; 53090 MW; 186FBA9CC5D745D CRC64;

Query Match 77.3%; Score 34; DB 2; Length 504;
Best Local Similarity 71.4%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10
Db 119 CVPRTGC 125

RESULT 26

```



OS4N64\_DICDI  
ID OS4N64\_DICDI PRELIMINARY; PRT; 507 AA.  
AC OS4N64;  
DT 13-SEP-2005 (T-EMBLrel. 31, Created)  
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)  
DE Hypothetical protein.  
GN ORFNames=DD80218721;  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota, Mycetozoa, Dictyostelidia, Dictyostelium.  
OX NCBI\_TaxId=44689;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=AX4;  
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
RA Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,  
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,  
RA Banker A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,  
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
RA Kerhrouou A., Nie X., Hall N., Anjard C., Hemphill L., Baon N.,  
RA Fairbrother P., Desany B., Just B., Morio T., Roest R., Churcher C.,  
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,  
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,  
RA Wadrop A., Felder M., Thangavelu M., Johnson D., Knights A.,  
RA Loulès H., Mangall K., Oliver C., Price C., Quail M.A.,  
RA Uruñihera H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,  
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,  
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzler M., Kay R.R.,  
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kapa A.;  
RT "The genome of the social amoeba Dictyostelium discoideum.";  
RL Nature 0:0-0(2005).  
-I- CAUTION: The sequence shown here is derived from an  
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
preliminary data.  
CC EMBL: AAF10100124; EMBL64710.1; -; Genomic\_DNA.  
DR Hypothetical protein.  
KW SEQUENCE 507 AA; 53214 MW; F5AFAC431BACEF5 CRC64;  
SQ  
Query Match 77.3%; Score 34; DB 2; Length 507;  
Best Local Similarity 71.4%; Pred. No. 2.1e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CXPXTC 10  
DB 184 CNPLTGC 190  
RESULT 27  
ID OS4N64\_TETNG PRELIMINARY; PRT; 670 AA.  
AC OS4N64;  
DT 13-SEP-2005 (T-EMBLrel. 31, Created)  
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)  
DE Hypothetical protein.  
GN ORFNames=GSTENG0019142001;  
OS Tetraodon nigroviridis (Green puffer).  
OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;  
OC Actinopterygii, Neopterygii, Teleostei, Euteleostei, Neoteleostei;  
OC Acanthomorpha, Acanthopterygii, Percormorpha, Tetraodontiformes;  
OC Tetraodontidae, Tetraodontidae; Tetraodon.  
OX NCBI\_TaxId=99883;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,  
RA Mauceli E., Bonneau L., Fischer C., Oufouf-Costat C., Bernot A.,  
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
RA Dassiva C., Saitanoubat M., Levy M., Boudet N., Castellano S.,  
RA Anthouard V., Uubin C., Castell V., Katinka M., Vacherie B.,

RA Biemont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,  
RA Cruaud C., Duprat S., Broctier P., Coutanceau J.P., Gonzy J.,  
RA Parra G., Lardier G., Chaple C., McKernan K.J., McEwan P., Bosak S.,  
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
RA Laudet V., Schachter V., Queller F., Sautin W., Scarpelli C.,  
RA Winkler P., Lander E.S., Weissbach J., Roest Croillins H.;  
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
the early vertebrate proto-karyotype.";  
RL Nature 431:946-957(2004).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC Genoscope; Whitehead Institute Centre for Genome Research;  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -I- CAUTION: The sequence shown here is derived from an  
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
preliminary data.  
CC EMBL: CAAB01014604; CAG00643.1; -; Genomic\_DNA.  
DR EMBL: CAAB01014604;  
FT NON TER 1  
SQ SEQUENCE 670 AA; 76484 MW; 0715D5B56262C30 CRC64;  
Query Match 77.3%; Score 34; DB 2; Length 670;  
Best Local Similarity 71.4%; Pred. No. 2.8e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CXPXTC 10  
DB 114 CXPXTC 120  
RESULT 28  
ID OSB3W6\_EMBNI PRELIMINARY; PRT; 870 AA.  
AC OSB3W6;  
DT 10-MAY-2005 (T-EMBLrel. 30, Created)  
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)  
DE Hypothetical protein.  
GN ORFNames=AN4764.2;  
OS Aspergillus nidulans FGSC A4.  
OC Eukaryota, Fungi, Ascomycota, Pezizomycotina, Eurotiomycetes;  
OC Eurotiales, Trichocomaceae; Emeritella.  
OX NCBI\_TaxId=227321;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=FGSC A4;  
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,  
RA Arachchi H.M., Barina N., Bastien V., Bloom T., Boguslavsky L.,  
RA Bouckgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,  
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,  
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,  
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,  
RA Garayna S., Gierke S., Graham L., Grand-Pierre N., Halez N.,  
RA Hagopian D., Hagoos B., Hall J., Horton L., Hulne W., Iliev I.,  
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,  
RA Kellis C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,  
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,  
RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Menes L.,  
RA Mihova T., Mienga V., Murphy T., Naylor T., Nguyen C., Nicol R.,  
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,  
RA Oliver J., Peterson K., Phunhthang P., Pletre N., Purcell S.,  
RA Rappaport A., Ramsamy U., Raymond C., Retta R., Rise C., Rogov P.,  
RA Roman J., Schuer B., Schupbach R., Sedman S., Severly P., Smirnov S.,  
RA Smith C., Spencer B., Strange-Thomann N., Stojanovic N., Stubbs M.,  
RA Talamas J., Teste S., Theodore O., Toppan K., Travers M.,  
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,  
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,  
RA Lander E.;  
RT "Genome Sequence of Aspergillus nidulans.";  
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
CC -I- CAUTION: The sequence shown here is derived from an  
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
preliminary data.

DR EMBL: AACD01000080: EAA60806.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 870 AA; 96093 MW; 2679DA01B40B338 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 870;  
 Best Local Similarity 71.4%; Pred. No. 3.5e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
 | | | | |  
 DB 478 CXPXTGC 484

RESULT 29

OS4N02 D1CD1 PRELIMINARY; PRT; 932 AA.  
 ID 054N02; D1CD1 PRELIMINARY;  
 AC 054N02;  
 DT 13-SEP-2005 (TREMBlrel. 31, Created)  
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)  
 DE Hypothetical protein.  
 OS ORFNames=DDB0186586;  
 GN Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=AX4;  
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
 RA Sugang R., Bertrman M., Song J., Olsen R., Szafianski K., Xu Q.,  
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,  
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gautier P., Fey P.,  
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Baeson N.,  
 RA Farbrother P., Deasny B., Just E., Morio T., Rost R., Churcher C.,  
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
 RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,  
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,  
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,  
 RA Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,  
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,  
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,  
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
 RA Williams J., Dear P.H., Noegel A.A., Bartell B., Kupa A.;  
 RA "The genome of the social amoeba Dictyostelium discoideum.";  
 RL Nature 0:0-0(2005).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC preliminary data.  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 DR EMBL: AAF101000124; EAL64630.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 932 AA; 99556 MW; EF21C56D1CCA5E22 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 932;  
 Best Local Similarity 71.4%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
 | | | | |  
 DB 609 CXPXTGC 615

RESULT 30  
 OS69K4 D1CD1 PRELIMINARY; PRT; 998 AA.  
 ID 0869K4; D1CD1 PRELIMINARY;  
 AC 0869K4;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)  
 DE Similar to Mus musculus (Mouse). Tenascin X (Hypothetical  
 protein).

GN ORFNames=DDB0167241;  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=AX4;  
 RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
 RA Sugang R., Bertrman M., Song J., Olsen R., Szafianski K., Xu Q.,  
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,  
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gautier P., Fey P.,  
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Baeson N.,  
 RA Farbrother P., Deasny B., Just E., Morio T., Rost R., Churcher C.,  
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
 RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,  
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,  
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,  
 RA Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,  
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,  
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,  
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
 RA Williams J., Dear P.H., Noegel A.A., Bartell B., Kupa A.;  
 RA "The genome of the social amoeba Dictyostelium discoideum.";  
 RL Nature 0:0-0(2005).  
 CC EMBL: AAF101000130; EAL69580.1; -; Genomic\_DNA.  
 DR EMBL: AAF101000130; EAL69580.1; -; Genomic\_DNA.  
 DR HSSP: P01132; IEGF.  
 DR GO: GO:0008305; C:integrin complex; IEA.  
 DR GO: GO:0005198; F:structural molecule activity; IEA.  
 DR GO: GO:0007160; P:cell-matrix adhesion; IEA.  
 DR InterPro: IPR000742; EGF\_2.  
 DR InterPro: IPR006209; EGF-like.  
 DR InterPro: IPR006210; IEGF.  
 DR InterPro: IPR000413; Integrin alpha.  
 DR InterPro: IPR002049; Laminin\_EGF.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF01839; FG-GAP; 1.  
 DR PRINTS: PRO0011; EGF\_1.  
 DR SMART: SM00181; EGF\_9.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_10.  
 DR PROSITE: PS01186; EGF\_2; 7.  
 DR PROSITE: PS00026; EGF\_3; 8.  
 KW Hypothetical protein.  
 SQ SEQUENCE 998 AA; 106001 MW; F79BFF394D3E2369 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 998;  
 Best Local Similarity 71.4%; Pred. No. 4e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
 | | | | |  
 DB 609 CXPXTGC 615

RESULT 31

OS5C92 DICI  
ID OS5C92 DICI PRELIMINARY; PRT; 999 AA.  
AC OS5C92;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DE Hypothetical protein.  
GN ORFNames=DD80216748;  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=AX4;  
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,  
RA Sugeng R., Berriman M., Song J., Olsen R., Szatanki K., Xu Q.,  
RA Tunggal B., Kummerfeld S., Madera M., Kontorov B.A., Rivero F.,  
RA Bankler A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,  
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
RA Kerhrou A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,  
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,  
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,  
RA Hauser H., James K., Quiles M., Mohan M.B., Salto T., Buchrieser C.,  
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,  
RA Loulès H., Mungall K., Oliver K., Price C., Quail M.A.,  
RA Urushihara H., Hernandez J., Rabinowitz E., Steffen D., Sanders M.,  
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,  
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
RA Williams J., Dear P.H., Noegel A.A., Bartell B., Kuspa A.;  
RT "The genome of the social amoeba Dictyostelium discoideum";  
RL Nature 0:0-0(2005).  
CC -!- CAUTION: The sequence shown here is derived from an  
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
preliminary data.  
CC EMBL; AAF10100011; EMBL72871.1; -; Genomic\_DNA.  
CD DR EMBL; AAF10100011; EMBL72871.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 999 AA; 106586 MW; FC387975B85D9C CRC64;

Query Match 77.3%; Score 34; DB 2; Length 999;  
Best Local Similarity 71.4%; Pred. No. 4e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTC 10  
DB 609 CNPLTGC 615

RESULT 32  
ID OS7XW5\_9TRYP PRELIMINARY; PRT; 1147 AA.  
AC OS7XW5;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DE Hypothetical protein.  
GN ORFNames=TP927.3.1320;  
OS Trypanosoma brucei.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxID=5691;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=GUTat10.1;  
RA Ghedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,  
RA Shallom J., Hou L., Djikeng A., Feldlyum T., Hostetler J.,  
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,  
RA Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,  
RA Wanless D., White O., Wortman J., Frazer C.W., El-Sayed N.M.A.;  
RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC159411; AX69554.1; -; Genomic\_DNA.  
KW Hypothetical protein.

SQ SEQUENCE 1147 AA; 124779 MW; 60C07617752B4AA7 CRC64;  
Query Match 77.3%; Score 34; DB 2; Length 1147;  
Best Local Similarity 71.4%; Pred. No. 4.6e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTC 10  
DB 343 CVPRTGC 349

RESULT 33  
ID Q6Y625\_9VIRU PRELIMINARY; PRT; 2217 AA.  
AC Q6Y625;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DE Polymerase.  
OS Lassa virus.  
OC Viruses; ssRNA negative-strand viruses; Arenaviridae; Arenavirus;  
OC Old world arenaviruses.  
OX NCBI\_TaxID=11620;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=CSF;  
RA Viet H.S., Gunther S.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY179174; AA059515.1; -; Genomic\_RNA.  
DR GO; GO:0003968; R.RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0019079; P:viral genome replication; IEA.  
DR InterPro; IPR010453; Arena\_RNA\_pol  
DR InterPro; IPR001000; Glyco\_Hydro\_10.  
DR InterPro; IPR007099; RNA\_pol\_NSVir.  
DR Pfam; PF06317; Arena\_RNA\_pol; 1.  
DR PROSITE; PS00591; GLYCOSYL\_HYDROL\_F10; UNKNOWN 1.  
SQ SEQUENCE 2217 AA; 253347 MW; 577DBF9A5A546EC7 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 2217;  
Best Local Similarity 71.4%; Pred. No. 8.5e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTC 10  
DB 1001 CGPMTGC 1007

RESULT 34  
ID Q4IR98\_GIBZE PRELIMINARY; PRT; 77 AA.  
AC Q4IR98;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DE Hypothetical protein.  
GN ORFNames=FG00260.1;  
OS Gibberella zeae PH-1.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.  
OX NCBI\_TaxID=229533;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=PH-1;  
RA Birren B., Nussbaum C., Abouelleil A., Allen N., Anderson S.,  
RA Arachchi H.M., Barina N., Bastien V., Bloom T., Boguslavsky L.,  
RA Boukhalter B., Butler J., Calvo S.E., Camarata J., Chang J.,  
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,  
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,  
RA Erickson U., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,  
RA Garayna S., Gierre S., Graham U., Grand-Pierre N., Hatz N.,  
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,  
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,

RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,  
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,  
RA Mathews C., Mauceli E., McCarthy M., Mejdrit J., Menes L.,  
RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,  
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,  
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,  
RA Rachtup A., Ramasamy U., Raymond C., Retta R., Riese C., Rogov P.,  
RA Roman J., Schauer S., Schupack R., Seaman S., Severy P., Smirnov S.,  
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,  
RA Talmay J., Tesfaye S., Theodore J., Topham K., Travers M.,  
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,  
RA Wu X., Wymann D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,  
RA Lander E.;  
RT "Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases."  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
CC EMBL; AACM0100010; EAA69670.1; -; Genomic\_DNA.  
KM Hypothetical protein.  
SQ SEQUENCE 77 AA; 7980 MW; 72702C4A47662E61 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 77;  
Best Local Similarity 71.4%; Pred. No. 57;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 51 CXPXTGC 57

RESULT 35  
067W06\_ORYZA PRELIMINARY; PRT; 269 AA.

AC 067W06; 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Putative senescence-associated protein 5.  
GN Name=OSUNBA0085J13.14;  
OS Oryza sativa (japonica cultivar-group). Embryophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaeae; Oryza.  
NCBI\_TaxID=39947;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC  
RT clone:OSUNBA0085J13.1";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB003565; BMD37413.1; -; Genomic\_DNA.  
DR Gramene; 067W06; -;  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR InterPro; IPR000301; Transmem\_4.  
DR Pfam; PF00335; Tetraspanin; I.  
DR PRINTS; PR00259; TMFOUR.  
DR PROSITE; PS00421; TMA\_1; 1.  
SQ SEQUENCE 269 AA; 25488 MW; 8B8EAF766DEB13 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 269;  
Best Local Similarity 71.4%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 172 CXPXTGC 178

RESULT 36  
OSNVV4\_9ARCH PRELIMINARY; PRT; 330 AA.  
ID OSNVV4\_9ARCH PRELIMINARY;  
AC OSNVV4;

DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
DE Predicted lipase synthase.  
GN ORFName=orf20;  
OS uncultured archaeon.  
OC Archaea; environmental samples.  
NCBI\_TaxID=115547;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Erkel C., Kemnitz D., Kube M., Rieke P., Chin K.-J., Dedysh S.,  
RA Reinhardt R., Conrad R., Liesack W.;  
RT "Retrieval of first genome data for rice cluster I methanogens by a  
RT combination of cultivation and molecular techniques";  
RL FEMS Microbiol. Ecol. 53:187-204(2005).  
DR EMBL; CR62685; CAH04852.1; -; Genomic\_DNA.  
DR GO; GO:0003824; F: catalytic activity; IEA.  
DR GO; GO:0005506; F: iron ion binding; IEA.  
DR InterPro; IPR006638; EIP3/MAB/Ni1B.  
DR InterPro; IPR007197; Radical SAM.  
DR Pfam; PF04055; Radical SAM.  
DR SMART; SM00729; EIP3; I.  
SQ SEQUENCE 330 AA; 35990 MW; 5090ECC290E75C34 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 330;  
Best Local Similarity 71.4%; Pred. No. 2.2e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 316 CXPXTGC 322

RESULT 37  
07XV21\_ORYZA PRELIMINARY; PRT; 464 AA.

AC 07XV21; 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE OSUNBA0064H22.11 protein.  
GN Name=OSUNBA0064H22.11;  
OS Oryza sativa (japonica cultivar-group). Embryophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaeae; Oryza.  
NCBI\_TaxID=39947;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA PubMed=12447439; DOI=10.1038/nature01183;  
RX Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,  
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,  
RA Meng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,  
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,  
RA Wu W., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,  
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Yu Y., Jia J., Zhang Y.,  
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,  
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,  
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang U., Li J., Hong G., Xue Y.,  
RA Han B.;  
RT "Sequence and analysis of rice chromosome 4.";  
RL Nature 420:316-320(2002).  
DR EMBL; AL066448; CAD40872.2; -; Genomic\_DNA.  
DR Gramene; 07XV21; -;

DR GO; GO:0004194; F: pepsin A activity; IEA.  
DR GO; GO:0006508; F: proteolysis and peptidolysis; IEA.  
DR InterPro; IPR003439; ABC\_transp\_like.  
DR InterPro; IPR001461; Peptidase\_A1.  
DR Pfam; PF00026; Asp; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; UNKNOWN 1.  
SQ SEQUENCE 464 AA; 45354 MW; 43B1443D3B732501 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 464;

Best Local Similarity 71.4%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 117 CDPCTGC 123

RESULT 38  
Q9LEZ5\_ARATH

ID Q9LEZ5\_ARATH PRELIMINARY; PRT; 1058 AA.

AC Q9LEZ5;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DE Hypothetical protein T22D6\_50.

GN Name=T22D6\_50;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC euroside II; Brassicales; Brassicaceae; Arabidopsis.

OC NCBI\_TaxID=3702;

OX NCBI\_TaxID=3702;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Bevan M., Terry N., Ardiles W., Buyschaert C., Dasseville R.,

RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,

RA Villarroel R., Gielen J., Van Montagu M., Bancroft I., Mewes H.W.,

RA Rudd S., Lemcke K., Mayer K.F.X.;

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RA EU Arabidopsis sequencing project;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL357612; CAB93712.1; -; Genomic\_DNA.

DR PIR; T50496; T50496.

DR GO; GO:0005324; F:ATP binding; IEA.

DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0006464; P:protein modification; IEA.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR01545; DEAD/DEAH\_N.

DR InterPro; IPR001650; Helicase\_C.

DR InterPro; IPR000626; Ubiqutin.

DR Pfam; PF00270; DEAD; 1.

DR Pfam; PF00271; Helicase\_C; 1.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00490; HELIC\_C; 1.

DR PROSITE; PSS0053; UBIQUITIN\_2; 1.

DR Hypothetical protein.

SQ SEQUENCE 1058 AA; 116855 MW; D95C267BF2672061 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 1058;

Best Local Similarity 71.4%; Pred. No. 6.6e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 1019 CDPCTGC 1025

RESULT 39

Q80TF6\_MOUSE

ID Q80TF6\_MOUSE PRELIMINARY; PRT; 1468 AA.

AC Q80TF6;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DR MKIA1300 protein (Fragment).

GN Name=Stard9; Synonyms=4831403C07R1k, MKIA1300;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC Tissue=Brain;

RA MEDLINE=22579291; PubMed=12693553;

RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,

RA Nakajima D., Nagase T., Ohara O., Koga H.;

RT "Prediction of the coding sequences of mouse homologues of KIAA gene:

RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous

RT cDNAs identified by screening of terminal sequences of cDNA clones

RT randomly sampled from size-fractionated libraries."

RL DNA Ref. 10:35-48(2003).

DR EMBL; AK122489; BAC65771.1; -; mRNA.

DR Ensembl; ENSMUSG00000033705; Mus musculus.

DR MGI; MGI:2444446; 4831403C07R1k.

DR MGI; MGI:3045258; Stard9.

DR InterPro; IPR002913; START\_lipid\_bd.

DR PROSITE; PSS0848; START; 1.

FT NON\_TER

SQ SEQUENCE 1468 AA; 161136 MW; 0431A231D506374 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 1468;

Best Local Similarity 71.4%; Pred. No. 9e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 387 CDPCTGC 393

RESULT 40

Q9ZU25\_RHIME

ID Q9ZU25\_RHIME PRELIMINARY; PRT; 64 AA.

AC Q9ZU25;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DR Hypothetical protein SMD21688.

GN OrderedlocusNames=RB1318; ORFNames=SMD21688;

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.

OC NCBI\_TaxID=382;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=1021;

RA MEDLINE=21396508; PubMed=11481431; DOI=10.1073/pnas.161294698;

RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,

RA Vorholter F.J., Hernandez-Lucas I., Becker A., Gouzy J.,

RA Golding B., Pehler A.;

RT "The complete sequence of the 1,683-kb, pSymB megaplasmid from the N2-

RT fixing endosymbiont Sinorhizobium meliloti."

RT Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).

DR EMBL; AL591985; CAC49718.1; -; Genomic\_DNA.

DR PIR; P96006; P96006.

KW Complete proteome; Hypothetical protein; Plasmid.

SQ SEQUENCE 64 AA; 7064 MW; A12006E21B8774D3 CRC64;

Query Match 72.7%; Score 32; DB 2; Length 64;

Best Local Similarity 57.1%; Pred. No. 75;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 15 CAPSSGC 21

RESULT 41

Q926M4\_LISIN

ID Q926M4\_LISIN PRELIMINARY; PRT; 123 AA.

AC Q926M4;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

```
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
PI P10035 protein.
Ordered locusNames=pl10035;
OS Listeria innocua.
OC plasmid pl1100.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
NC NCB1_TaxID=1642;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669; DOI=10.1126/science.1063447;
RA Glaeser P., Frangoul L., Buchrieser C., Ruemik C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Chabrie A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussutget O.,
RA Entian K.-D., Fshih H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
RA Jones L.-M., Kaerst U., Klett U., Kuhn M., Kunst F., Kurapkut G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Norisler G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tleier A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
"Comparative genomics of Listeria species.";
Science 294:849-852(2001).
RL EMBL; AL592102; CAC42033.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016564; F:transcriptional repressor activity; IEA.
DR GO; GO:0016481; P:negative regulation of transcription; IEA.
DR GO; GO:0046685; P:response to arsenic; IEA.
DR InterPro; IPR010712; Arsd.
DR InterPro; IPR012335; ThioRedoxin_fold.
DR Pfam; PF06953; Arsd; 1.
KW Complete proteome; Plasmid.
SQ SEQUENCE 123 AA; 13164 MW; 5AB2979019CFF14F CRC64;

Query Match 72.7%; Score 32; DB 2; Length 123;
Best Local Similarity 57.1%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | |
Db 116 CSPSSGC 122

RESULT 42
OSTN00 ANOGA PRELIMINARY; PRT; 222 AA.
AC OSTN00.
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
DE ENSANGP00000028786 (Fragment).
ORFNames=ENSANG0000002818;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCB1_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST.
RG The Anopheles gambiae Sequence Committee;
RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST.
RG The Anopheles gambiae Sequence Committee;
RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
```

```
CC preliminary data.
DR EMBL; AAB01008986; EAL38762.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 222 AA; 21986 MW; 3628579D88CCE889 CRC64;

Query Match 72.7%; Score 32; DB 2; Length 222;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | |
Db 78 CSPSSGC 84

RESULT 43
OSTM29 ANOGA PRELIMINARY; PRT; 231 AA.
AC OSTM29.
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
DE 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE ENSANGP00000026276 (Fragment).
ORFNames=ENSANG0000002818;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCB1_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST.
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RN Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST.
RG The Anopheles gambiae Sequence Committee;
RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAB01008986; EAL38763.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 231
SQ SEQUENCE 231 AA; 23375 MW; 888B95ACB094A91B CRC64;

Query Match 72.7%; Score 32; DB 2; Length 231;
Best Local Similarity 57.1%; Pred. No. 2.5e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10
| | | |
Db 86 CSPSSGC 92

RESULT 44
O6Z8K1 ORYSA PRELIMINARY; PRT; 520 AA.
AC O6Z8K1.
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Aspartyl protease-like.
GN Name=P0431B06.25;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCB1_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
```

```

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004774; BAB15642.1; -; Genomic_DNA.
DR HSSP; P56272; 1AMS.
DR Gramene; O628K1; -.
DR GO; GO:0004194; F:pepsin A activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001461; Peptidase A1.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00026; Asp_1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
KW Protease.
SQ SEQUENCE 520 AA; 56271 MW; 89468ACBA8C0CA1C CRC64;

Query Match
Best Local Similarity 72.7%; Score 32; DB 2; Length 520;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10
Db 170 CSPASGC 176

RESULT 45
O9F126 ARATH PRELIMINARY; PRT; 1152 AA.
AC O9F126_
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MDN11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93937451; PubMed=10470850;
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
RT Sequence features of the regions of 1,011,550 bp covered by seventeen
RT P1 and TAC clones."
RL DNA Res. 6:183-195 (1999).
DR EMBL; AB017064; BAB1076.1; -; Genomic_DNA.
DR EMBL; AB023039; BAB1076.1; JOINED; Genomic_DNA.
DR GO; GO:0005488; P:binding; IEA.
DR InterPro; IPR011989; ARM-like.
SQ SEQUENCE 1152 AA; 126806 MW; FA379C6BP94CA04 CRC64;

Query Match
Best Local Similarity 72.7%; Score 32; DB 2; Length 1152;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10
Db 475 CSPXTGC 481

RESULT 46
O9WU17 MESAU PRELIMINARY; PRT; 162 AA.
ID O9WU17_
AC O9WU17_
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Receptor tyrosine kinase (Fragment).
GN Name:TYKc.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

```

```

OC Muridae; Cricetinae; Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hansen H.T., Lehman M.N., Stevens P.J.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF104241; MAD21012.1; -; mRNA.
DR HSSP; O62838; 1LUF.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. . .; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002011; ReceptTyKinseII.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR ProDom; PD000001; Prot_kinase_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
KW Kinase; Receptor.
FT NON_TER 1
FT TER 162
SQ SEQUENCE 162 AA; 18209 MW; 29DF9E933BECB243 CRC64;

Query Match
Best Local Similarity 70.5%; Score 31; DB 2; Length 162;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10
Db 96 CSPSAGC 102

RESULT 47
O8T415 DROME PRELIMINARY; PRT; 166 AA.
ID O8T415_
AC O8T415_
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
DE AT04449p (CG31128-PA).
GN ORFName=CG31128;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Stapleton M., Broksstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Padleb J., Paragias V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Bendandi J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Broksstein P., Brothier P.,

```



RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,  
RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I.C., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Flitschmann W.,  
RA Foster C., Gabrielian A.B., Gary N.S., Gelbart W.M., Glaeser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mikhina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svrtskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
[3]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svrtskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
RT *melanogaster* euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
[4]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svrtskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
RT a genomic perspective.";  
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
[5]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,  
RA Smith C.D., Tuzy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
[6]  
RP NUCLEOTIDE SEQUENCE.  
RG Berkeley *Drosophila* Genome Project;  
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svrtskas R., Smith E.,  
RA Xu C., Rubin G.;  
RT "Drosophila melanogaster release 4 sequence."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[7]  
RP NUCLEOTIDE SEQUENCE.  
RG FlyBase;

RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY084097; AA189835.1; -; mRNA.  
DR EMBL; AE003748; AA13997.1; -; Genomic\_DNA.  
DR Ensembl; CG31128; *Drosophila melanogaster*.  
DR FlyBase; FBgn0051128; CG31128.  
SQ SEQUENCE 166 AA; 18434 MW; 054DC42BDE653B76 CRC64;  
Query Match 70.5%; Score 31; DB 2; Length 166;  
Best Local Similarity 57.1%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 4 CKPXTGC 10  
DB 142 CSPTAGC 148  
RESULT 48  
ID 04H68\_9DEIO PRELIMINARY; PRT; 183 AA.  
AC 04H68;  
DT 13-SEP-2005 (TREMBLrel. 31, Created)  
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)  
DE Hypothetical protein.  
GN ORFNames=DgeODRAFT\_2849;  
OS Deinococcus geothermalis DSM 11300.  
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;  
OC Deinococcaceae; Deinococcus.  
OX NCBI\_TaxID=319795;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=DSM 11300;  
RG US DOE Joint Genome Institute (JGI-PGP);  
RA Copeland A., Lucas S., Lapidus A., Barry K., Deter C., Glavina T.,  
RA Hammon N., Istrati S., Pittluck S., Richardson P.;  
RT "Sequencing of the draft genome assembly of *Deinococcus geothermalis*  
RT DSM 11300.";  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
RN (2)  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=DSM 11300;  
RG US DOE Joint Genome Institute (JGI-ORNL);  
RA Lartier F., Land M.;  
RT "Annotation of the draft genome assembly of *Deinococcus geothermalis*  
RT DSM 11300.";  
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
CC DR EMBL; AAHE0100014; EAL61853.1; -; Genomic\_DNA.  
CC KM Hypothetical protein.  
SQ SEQUENCE 183 AA; 19839 MW; C5F6864281158966 CRC64;  
Query Match 70.5%; Score 31; DB 2; Length 183;  
Best Local Similarity 57.1%; Pred. No. 3.1e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CKPXTGC 10  
DB 167 CSPTAGC 173  
RESULT 49  
ID 04S7H4\_TETNG PRELIMINARY; PRT; 229 AA.  
AC 04S7H4;  
DT 13-SEP-2005 (TREMBLrel. 31, Created)  
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)  
DE Chromosome 13 SCAR14715, whole genome shotgun sequence.  
GN ORFNames=GSTENG00022782001;  
OS Tetraodon nigroviridis (green puffer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 : Search time 13.4522 Seconds  
(without alignments)  
92.983 Million cell updates/sec

Title: US-09-932-322-1

Perfect score: 44  
Sequence: 1 XXXCKPXTGCKXX 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 1000 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	81.8	601	2 A27020	DIF-induced preste
2	34	77.3	145	2 S07983	phospholipase A2 h
3	34	77.3	312	2 T05348	hypothetical prote
4	33	75.0	1058	2 T50496	hypothetical prote
5	32	72.7	64	2 P96006	hypothetical prote
6	31	70.5	482	2 T48397	S-receptor kinase-
7	31	70.5	643	2 T25473	hypothetical prote
8	30	68.2	129	2 A24255	chortion class A pr
9	30	68.2	132	2 B24255	chortion class A pr
10	30	68.2	151	2 AB0331	sigma B factor reg
11	30	68.2	199	2 T49486	hypothetical prote
12	30	68.2	200	2 S04926	wound-induced prote
13	30	68.2	201	2 T07729	wound-induced prote
14	30	68.2	211	2 T04098	CBP20 preproprotei
15	30	68.2	211	2 S04927	wound-induced prote
16	30	68.2	222	1 MWVZB4	antithrombin-III h
17	30	68.2	224	2 E84326	hypothetical prote
18	30	68.2	336	2 A83801	ABC transporter (s
19	30	68.2	345	1 MWVZM2	antithrombin-III h
20	30	68.2	406	2 T23898	hypothetical prote
21	30	68.2	430	2 T23899	hypothetical prote
22	30	68.2	551	2 S01793	arylsulfatase (EC
23	30	68.2	551	2 S07089	arylsulfatase (EC
24	30	68.2	648	2 T23864	hypothetical prote
25	30	68.2	1722	2 E89753	protein P11C7.4 li
26	29	65.9	46	2 H71262	hypothetical prote
27	29	65.9	119	2 B45937	early chortion prote
28	29	65.9	119	2 S24292	chortion protein -
29	29	65.9	119	2 S24294	chortion protein -

30	29	65.9	119	2 S24291	Chortion protein -
31	29	65.9	121	2 S24293	Chortion class CA p
32	29	65.9	147	2 E82523	hypothetical prote
33	29	65.9	163	2 T33130	hypothetical prote
34	29	65.9	245	2 F84680	hypothetical prote
35	29	65.9	333	2 P90172	hypothetical prote
36	29	65.9	348	2 T35248	probable oxidoredu
37	29	65.9	375	2 A83636	hypothetical prote
38	29	65.9	379	2 S14885	hypothetical prote
39	29	65.9	449	2 E86676	hypothetical prote
40	29	65.9	507	2 T23375	hypothetical prote
41	29	65.9	513	2 D88991	protein apx-1 (imp
42	29	65.9	685	2 JC7570	Delta-4 protein -
43	29	65.9	686	2 JC7569	hypothetical prote
44	29	65.9	753	2 T28787	hypothetical prote
45	29	65.9	843	1 A27340	complement C7 prec
46	29	65.9	1101	2 T16840	hypothetical prote
47	29	65.9	1146	2 T02766	adhesin Wt-1 - Aje
48	29	65.9	1743	2 T26859	hypothetical prote
49	29	65.9	3635	2 T10053	laminin alpha 5 ch
50	28	63.6	74	2 S25773	testis-specific pr
51	28	63.6	95	2 A13171	conserved hypothet
52	28	63.6	103	2 G84741	hypothetical prote
53	28	63.6	105	2 S23061	hypothetical prote
54	28	63.6	110	2 S16496	Chortion protein -
55	28	63.6	154	2 G72564	hypothetical prote
56	28	63.6	166	2 B44938	cysteine proteinas
57	28	63.6	178	2 S76878	hypothetical prote
58	28	63.6	207	2 G69381	flagellin (flaB1-2
59	28	63.6	207	2 B71943	hypothetical prote
60	28	63.6	218	2 T33545	hypothetical prote
61	28	63.6	237	2 I47031	insulin-like growt
62	28	63.6	248	2 D66012	insulin-like growt
63	28	63.6	248	2 D91166	probable membrane
64	28	63.6	254	2 I46603	insulin-like growt
65	28	63.6	254	2 I46599	insulin-like growt
66	28	63.6	254	2 JC1464	insulin-like growt
67	28	63.6	258	1 B37252	insulin-like growt
68	28	63.6	258	2 A45403	insulin-like growt
69	28	63.6	316	2 AG0121	probable sideropho
70	28	63.6	331	2 AC0842	glycine betaine-bi
71	28	63.6	340	1 WMBEL1	latency-related pr
72	28	63.6	449	1 L1PG	retinoic acid rece
73	28	63.6	453	2 I50674	ankyrin-related pr
74	28	63.6	491	2 B87452	arylsulfatase (EC
75	28	63.6	567	2 A37362	urease (EC 3.5.1.5
76	28	63.6	569	2 T50711	hypothetical prote
77	28	63.6	628	2 S50524	acetic acid glyco
78	28	63.6	744	2 A43353	protein P9C16.17 l
79	28	63.6	873	2 H96503	antigen receptor
80	28	63.6	910	2 A34721	antigen receptor
81	28	63.6	911	2 B34721	antigen receptor
82	28	63.6	919	2 A3248	antigen receptor
83	28	63.6	980	2 T00045	cellulodextrin phosp
84	28	63.6	1035	2 S18512	cell division cont
85	28	63.6	1172	2 P96503	protein P9C16.13 l
86	28	63.6	1220	2 A56136	jaaged protein pre
87	28	63.6	1332	2 T23024	hypothetical prote
88	28	63.6	1522	2 H88380	protein T22F7.3 li
89	28	63.6	1620	2 T27283	hypothetical prote
90	28	63.6	1717	1 A45558	epidermal growth f
91	28	63.6	2395	1 S50820	surface protein ty
92	28	63.6	2533	2 T28675	alpha-51D immobili
93	28	63.6	2533	2 T28674	alpha-51D-immobili
94	28	63.6	2543	2 T31687	surface antigen - P
95	28	63.6	2704	2 S09118	G surface protein
96	28	63.6	2718	2 A23475	G surface protein
97	27	61.4	19	2 A43379	alpha-conotoxin SI
98	27	61.4	46	1 KECX	crabmin - Alysini
99	27	61.4	55	2 A40186	ribosomal protein
100	27	61.4	64	2 P82845	hypothetical prote
101	27	61.4	75	2 S17156	hypothetical prote
102	27	61.4	87	2 S41306	hypothetical prote

Chortion protein -  
Chortion class CA p  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
probable oxidoredu  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
protein apx-1 (imp  
Delta-4 protein -  
hypothetical prote  
complement C7 prec  
hypothetical prote  
adhesin Wt-1 - Aje  
hypothetical prote  
laminin alpha 5 ch  
testis-specific pr  
conserved hypothet  
hypothetical prote  
Chortion protein -  
hypothetical prote  
cysteine proteinas  
hypothetical prote  
flagellin (flaB1-2  
hypothetical prote  
hypothetical prote  
insulin-like growt  
insulin-like growt  
insulin-like growt  
insulin-like growt  
insulin-like growt  
probable membrane  
insulin-like growt  
insulin-like growt  
insulin-like growt  
insulin-like growt  
probable sideropho  
glycine betaine-bi  
latency-related pr  
retinoic acid rece  
ankyrin-related pr  
arylsulfatase (EC  
urease (EC 3.5.1.5  
hypothetical prote  
acetic acid glyco  
protein P9C16.17 l  
antigen receptor  
antigen receptor  
antigen receptor  
antigen receptor  
cellulodextrin phosp  
cell division cont  
protein P9C16.13 l  
jaaged protein pre  
hypothetical prote  
protein T22F7.3 li  
hypothetical prote  
epidermal growth f  
surface protein ty  
alpha-51D immobili  
alpha-51D-immobili  
surface antigen - P  
G surface protein  
G surface protein  
alpha-conotoxin SI  
crabmin - Alysini  
ribosomal protein  
hypothetical prote  
hypothetical prote  
hypothetical prote

103	27	61.4	89	2	T42019	Ser/Thr protein ph
104	27	61.4	99	2	D70582	probable pe protei
105	27	61.4	118	1	PSUT1B	phospholipase A2 (
106	27	61.4	118	2	S52549	thionin variant Th
107	27	61.4	120	1	W4WL42	B4 protein - human
108	27	61.4	120	2	E72490	hypothetical prote
109	27	61.4	121	2	F81973	probable ribonucle
110	27	61.4	121	2	H81027	ribonuclease P pro
111	27	61.4	123	2	T19391	hypothetical prote
112	27	61.4	125	2	S52548	thionin variant Th
113	27	61.4	125	2	S52550	thionin variant Th
114	27	61.4	128	2	A72678	hypothetical prote
115	27	61.4	130	2	T08584	hypothetical prote
116	27	61.4	134	2	AD2592	cytidine deaminase
117	27	61.4	135	2	HSX132	histone H3.2 - Afr
118	27	61.4	136	1	S52545	chionin variant Th
119	27	61.4	145	2	F97374	cytidine deaminase
120	27	61.4	150	2	T49563	hypothetical prote
121	27	61.4	155	2	A10545	hypothetical prote
122	27	61.4	173	2	T11112	NADH2 dehydrogenas
123	27	61.4	190	2	T05492	thiamatin homolog
124	27	61.4	219	2	T51074	hypothetical prote
125	27	61.4	220	2	H81434	superoxide dismuta
126	27	61.4	231	2	S55078	hypothetical prote
127	27	61.4	234	2	T36162	probable integral
128	27	61.4	237	2	S08073	cyclic nucleotide
129	27	61.4	237	2	T25152	hypothetical prote
130	27	61.4	239	2	A25317	gag polyprotein -
131	27	61.4	246	2	A64326	hypothetical prote
132	27	61.4	248	2	S30194	ribosomal protein
133	27	61.4	249	1	R3HU6	ribosomal protein
134	27	61.4	249	1	R3MS6	ribosomal protein
135	27	61.4	249	1	R3RMS6	ribosomal protein
136	27	61.4	249	2	UC4145	ribosomal protein
137	27	61.4	249	2	S41468	ribosomal protein
138	27	61.4	249	2	T50003	40S ribosomal prote
139	27	61.4	250	2	T05100	ribosomal protein
140	27	61.4	251	2	T04335	ribosomal protein
141	27	61.4	253	2	S09215	erythromycin resis
142	27	61.4	254	2	S35743	p19 protein - avia
143	27	61.4	255	2	I38426	lymphocyte activat
144	27	61.4	256	2	B32393	T-cell antigen 4-1
145	27	61.4	259	2	T21011	hypothetical prote
146	27	61.4	265	2	C87679	cyclo protein limp
147	27	61.4	270	2	B84578	probable senescenc
148	27	61.4	271	2	S12783	OX40 antigen precu
149	27	61.4	272	2	I48700	gene ox40 protein
150	27	61.4	277	2	A60771	B-cell activation
151	27	61.4	277	2	I37552	OX40 homolog - hum
152	27	61.4	305	2	A46476	B cell-associated
153	27	61.4	313	2	S44208	extracellular matr
154	27	61.4	318	2	T29479	hypothetical prote
155	27	61.4	321	2	S13550	another-specific pr
156	27	61.4	325	2	B43692	t2 protein - rabbl
157	27	61.4	326	1	GOVZML	probable transcrip
158	27	61.4	336	2	C95902	glycine betaine/pr
159	27	61.4	330	1	BLRCGP	hypothetical prote
160	27	61.4	330	2	T25169	hypothetical prote
161	27	61.4	330	2	A85916	glycine betaine/pr
162	27	61.4	330	2	F91071	glycine betaine-bi
163	27	61.4	334	2	AG0322	fructose-bisphosph
164	27	61.4	337	2	A10427	hypothetical prote
165	27	61.4	344	2	T17266	lipase homolog T29
166	27	61.4	381	2	T06696	probable myo-inosi
167	27	61.4	388	2	T36191	hypothetical prote
168	27	61.4	392	2	H84679	GTP-binding protei
169	27	61.4	418	2	F84326	exo-alpha-stalidas
170	27	61.4	432	2	TS1468	hypothetical prote
171	27	61.4	432	2	AE3341	cobg protein (EC 1
172	27	61.4	456	2	UC7328	amino acid transpo
173	27	61.4	487	2	S28267	pherophorin I prec
174	27	61.4	512	2	S28267	conserved hypothe
175	27	61.4	528	2	B75310	
176	27	61.4	533	2	H71492	probable hsp-60 -
177	27	61.4	534	2	B81654	60 kDa chaperonin
178	27	61.4	539	2	T31901	hypothetical prote
179	27	61.4	561	1	SYRPM	malate synthase (E
180	27	61.4	562	2	T48413	malate synthase-ii
181	27	61.4	563	2	C95874	conserved hypothet
182	27	61.4	566	2	T07690	malate synthase (E
183	27	61.4	566	2	S15387	malate synthase (E
184	27	61.4	567	1	SYCNMU	malate synthase (E
185	27	61.4	567	1	SYCSM2	malate synthase (E
186	27	61.4	568	1	SYKXMA	protein f54Cl.1 (f
187	27	61.4	581	2	B87768	probable acylamino
188	27	61.4	591	2	H72474	guanosine-3',5'-bi
189	27	61.4	607	2	F70124	gag polyprotein -
190	27	61.4	701	2	S35430	gag polyprotein -
191	27	61.4	720	1	A55160	Trg protein - frui
192	27	61.4	824	1	S50767	S-receptor kinase
193	27	61.4	826	1	Q0BE11	ribonucleoside-dip
194	27	61.4	887	2	S70642	ubiquitin ligase N
195	27	61.4	964	2	S51324	pullulanase - spin
196	27	61.4	987	2	D97029	ribonucleotide red
197	27	61.4	1046	2	A26838	prestalk protein p
198	27	61.4	3133	2	S52093	hemocytin - silkwo
199	27	61.4	3566	1	A40701	tenascin-X precurs
200	27	61.4	4351	2	T00252	MEGF1 protein - ra
201	27	61.4	4351	2	A26363	cardioactive pepti
202	27	61.4	9	2	S27233	potassium channel
203	27	61.4	35	2	S69599	hypothetical prote
204	27	61.4	45	2	D82788	hypothetical prote
205	27	61.4	53	2	E82628	hypothetical prote
206	27	61.4	56	1	WTF	testis-specific pr
207	27	61.4	63	2	S25772	somatotropin precu
208	27	61.4	77	2	PC1017	T-cell surface gly
209	27	61.4	80	2	C49050	hypothetical prote
210	27	61.4	89	2	S75008	transposase sb1392
211	27	61.4	93	2	S75008	phospholipase A2 (
212	27	61.4	118	1	PSL73E	phospholipase A2 (
213	27	61.4	118	2	C34860	phospholipase A2 (
214	27	61.4	118	2	F34860	phospholipase A2 h
215	27	61.4	119	1	PSNOA1	alpha-amyase inh
216	27	61.4	121	2	S16920	alpha-amyase inh
217	27	61.4	123	1	WMTA	hypothetical prote
218	27	61.4	123	2	G75336	hypothetical prote
219	27	61.4	126	1	B64025	conserved hypothet
220	27	61.4	129	2	A27562	T-cell activation
221	27	61.4	129	2	AB2728	hypothetical prote
222	27	61.4	134	2	A59055	phospholipase A2 (
223	27	61.4	134	2	B59055	phospholipase A2 (
224	27	61.4	134	2	E70882	hypothetical prote
225	27	61.4	145	2	S01390	phospholipase A2 (
226	27	61.4	145	2	I54504	NGF5 secreted spl
227	27	61.4	151	2	T06517	alpha-amyase inh
228	27	61.4	154	2	AB2509	hypothetical prote
229	27	61.4	160	2	S76328	hypothetical prote
230	27	61.4	162	1	PSHBA	phospholipase A2 (
231	27	61.4	170	2	E82145	conserved hypothet
232	27	61.4	171	2	S57894	laminin - Hydra vu
233	27	61.4	172	2	S72644	pilx protein precu
234	27	61.4	176	2	T02217	NBS-IRR type resis
235	27	61.4	181	2	T03720	4-coumarate-CoA 11
236	27	61.4	182	2	T02091	ribosomal protein
237	27	61.4	187	2	G70688	probable lipv prot
238	27	61.4	188	2	H38192	chitin synthase (E
239	27	61.4	195	2	C83076	type 4 fimbrial bl
240	27	61.4	198	2	A45067	laminin B1 chain v
241	27	61.4	199	2	UE0351	OX40 ligand protei
242	27	61.4	202	2	T48709	hypothetical prote
243	27	61.4	204	2	A54560	TPA-induced protei
244	27	61.4	205	1	P8XRBH	nonstructural prot
245	27	61.4	205	2	T04370	pern protein - ba
246	27	61.4	207	2	A64555	hypothetical prote
247	27	61.4	207	2	AC3555	hypothetical prote
248	27	61.4	212	2	A72648	

249	26	59.1	217	1	VCfMS	coat protein - alf	322	26	59.1	383	2	F86411	putroetin FlK23.16
250	26	59.1	218	1	VCfV2S	coat protein - alf	323	26	59.1	384	2	S68410	lipase Arab-1 - Ar
251	26	59.1	219	2	B83396	probable COA trans	324	26	59.1	387	2	B84721	hypothetical prote
252	26	59.1	220	2	S47166	superoxide dismuta	325	26	59.1	390	2	F69452	probable aspartate
253	26	59.1	221	1	VCfM42	coat protein - alf	326	26	59.1	390	2	B86411	protein FlK23.18 (
254	26	59.1	221	1	VCfMYS	coat protein - alf	327	26	59.1	390	2	G82844	cysteine synthase
255	26	59.1	225	2	T20325	hypothetical prote	328	26	59.1	393	2	T51108	dehydroatase [impor
256	26	59.1	228	2	F75390	hypothetical prote	329	26	59.1	394	2	A84672	probable lipase (I
257	26	59.1	229	2	S66342	ribonuclease II (B	330	26	59.1	397	2	T46223	hypothetical prote
258	26	59.1	230	2	S53506	ribonuclease LB (B	331	26	59.1	400	2	C75336	serine proteinase,
259	26	59.1	232	2	T33704	hypothetical prote	332	26	59.1	403	2	UE0114	zinc-finger protei
260	26	59.1	237	2	S53507	starvation-induced	333	26	59.1	404	2	T06761	hypothetical prote
261	26	59.1	237	2	T05973	permatin homolog P	334	26	59.1	407	2	AE3615	glycosyl transfera
262	26	59.1	247	2	AD3649	transcription regu	335	26	59.1	408	2	G70522	probable transposa
263	26	59.1	249	2	T35589	probable secreted	336	26	59.1	410	2	T47926	hypothetical protea
264	26	59.1	253	2	S73473	DNA polymerase III	337	26	59.1	410	2	T51212	related to integra
265	26	59.1	253	2	T25768	hypothetical prote	338	26	59.1	414	2	T38742	hypothetical prote
266	26	59.1	258	2	A29088	Smb class II histo	339	26	59.1	420	2	C69532	conserved hypotet
267	26	59.1	266	2	AG3000	1-acyl-sn-glycerol	340	26	59.1	422	2	D86339	hypothetical protea
268	26	59.1	266	2	B98283	hypothetical prote	341	26	59.1	423	2	A41207	collagen 13, nonfi
269	26	59.1	269	2	T04095	ribonuclease S hom	342	26	59.1	426	2	B84531	hypothetical prote
270	26	59.1	270	2	F69362	conserved hypotet	343	26	59.1	427	2	D84540	hypothetical prote
271	26	59.1	271	2	C86251	hypothetical prote	344	26	59.1	427	2	T48159	hypothetical prote
272	26	59.1	274	2	F85741	hypothetical prote	345	26	59.1	430	2	B64129	adenosylmethionine
273	26	59.1	276	2	C84471	probable U4/U6\smu	346	26	59.1	433	2	F84215	aminopeptidase [im
274	26	59.1	277	2	A75112	hypothetical prote	347	26	59.1	435	2	I54182	tumor necrosis fac
275	26	59.1	277	2	G71021	hypothetical prote	348	26	59.1	446	2	T47744	hypothetical prote
276	26	59.1	281	2	AD2755	conserved hypotet	349	26	59.1	446	2	C75087	hydroxymethylpyrim
277	26	59.1	281	2	AB3367	maze protein (Impo	350	26	59.1	446	2	B71057	probable thiamin b
278	26	59.1	284	2	E70506	hypothetical prote	351	26	59.1	455	2	I54353	tetracycline trans
279	26	59.1	285	2	T47500	hypothetical prote	352	26	59.1	463	2	A83762	RNA methyltransfer
280	26	59.1	285	2	E75507	hypothetical prote	353	26	59.1	464	2	S75831	hypothetical prote
281	26	59.1	285	2	H95969	hypothetical prote	354	26	59.1	466	2	F69806	RNA methyltransfer
282	26	59.1	286	2	B84807	probable RNA-bind	355	26	59.1	468	1	B66114	glycoprotein gp13
283	26	59.1	292	2	C97536	maze protein (Af22	356	26	59.1	468	1	VG8EBH	glycoprotein gp13
284	26	59.1	295	2	T35193	sporulation trans	357	26	59.1	481	2	T48404	3-deoxy-D-manno-oc
285	26	59.1	297	2	T18960	hypothetical prote	358	26	59.1	495	2	F43496	envelope protein E
286	26	59.1	297	2	S06267	surface antigen H	359	26	59.1	495	2	B43496	envelope protein E
287	26	59.1	298	2	C82318	chromosome initiat	360	26	59.1	495	2	C43496	envelope protein E
288	26	59.1	301	2	H70644	hypothetical prote	361	26	59.1	495	2	B43496	envelope protein E
289	26	59.1	305	2	A97158	stage III sporulat	362	26	59.1	495	2	G43496	envelope protein E
290	26	59.1	306	2	F64185	D-alanine-D-alalan	363	26	59.1	495	2	S06740	envelope protein E
291	26	59.1	306	2	H86695	ABC transporter pe	364	26	59.1	495	2	H43496	envelope protein E
292	26	59.1	308	2	T35835	hypothetical prote	365	26	59.1	495	2	S49137	genome polyprotein
293	26	59.1	310	2	F64146	hypothetical prote	366	26	59.1	495	2	S06741	envelope protein E
294	26	59.1	312	2	A61183	hypothetical prote	367	26	59.1	495	2	S11482	envelope glycoprot
295	26	59.1	314	2	AB0144	conserved hypotet	368	26	59.1	495	2	S06747	envelope protein E
296	26	59.1	315	2	T18956	hypothetical prote	369	26	59.1	496	2	S51668	tyrosine kinase -
297	26	59.1	319	2	AB0641	pseudouridylyate sy	370	26	59.1	499	2	S74224	aldehyde dehydroge
298	26	59.1	319	2	C64852	probable pseudouri	371	26	59.1	500	2	JC7668	depeptidyl-peptida
299	26	59.1	319	2	D85671	hypothetical prote	372	26	59.1	503	2	T05347	hypothetical prote
300	26	59.1	319	2	H90811	hypothetical prote	373	26	59.1	505	1	PIWLRB	Li protein - cotro
301	26	59.1	321	1	Q0EC15	hypothetical prote	374	26	59.1	510	1	PWQFPA	H+-transporting tw
302	26	59.1	321	2	F91104	hypothetical 34K p	375	26	59.1	510	1	D96741	hypothetical prote
303	26	59.1	321	2	A85950	hypothetical prote	376	26	59.1	511	1	S60287	glucose-6-phosphat
304	26	59.1	326	2	D72689	hypothetical prote	377	26	59.1	512	2	AG3527	D-xylose transport
305	26	59.1	328	2	E85842	probable superoxid	378	26	59.1	515	2	T52610	glucose-6-phosphat
306	26	59.1	332	2	B71165	hypothetical prote	379	26	59.1	519	2	T43756	ATPase subunit 1 (
307	26	59.1	336	2	AG2686	ABC transporter, s	380	26	59.1	522	2	S53842	H+-transporting tw
308	26	59.1	336	2	D97468	hypothetical prote	381	26	59.1	524	2	G73397	MG68 homolog Ro2_
309	26	59.1	338	2	E75576	oxidoreductase - D	382	26	59.1	528	2	S13641	H+-transporting tw
310	26	59.1	338	2	B96976	probable ABC trans	383	26	59.1	530	2	T30505	hypothetical prote
311	26	59.1	341	2	T33949	hypothetical prote	384	26	59.1	531	2	T43551	multitidng resistan
312	26	59.1	341	2	C86461	PI4M2.10 protein -	385	26	59.1	531	2	C83153	conserved hypotet
313	26	59.1	344	2	A75576	oxidoreductase - D	386	26	59.1	532	2	T52442	hypothetical prote
314	26	59.1	353	2	A87469	ubiquinol oxidase	387	26	59.1	545	1	A39827	4-connarate-CoA II
315	26	59.1	359	2	T20575	hypothetical prote	388	26	59.1	545	1	B39827	4-connarate-CoA II
316	26	59.1	360	2	T33835	hypothetical prote	389	26	59.1	551	2	C71606	ATP synthase alpha
317	26	59.1	364	1	PAQXF	fructose-bisphosph	390	26	59.1	554	2	T02445	probable U4/U6 sma
318	26	59.1	367	2	T48048	hypothetical prote	391	26	59.1	555	2	U01405	genome polyprotein
319	26	59.1	367	2	G85362	hypothetical prote	392	26	59.1	555	2	U01404	genome polyprotein
320	26	59.1	378	1	A70571	probable hexosyltr	393	26	59.1	559	2	T09135	hypothetical prote
321	26	59.1	381	2	T45940	hypothetical prote	394	26	59.1	559	2	B64556	para-aminobenzoate

395	26	59.1	560	2	A13216	hypotheical prote
396	26	59.1	561	2	C96837	nodulin-like prote
397	26	59.1	563	2	S33957	coat protein gamma
398	26	59.1	567	2	A71951	p-aminobenzoate sy
399	26	59.1	588	2	T49210	hypotheical prote
400	26	59.1	586	2	T35549	hypotheical prote
401	26	59.1	597	2	G84825	probable CCGH-type
402	26	59.1	603	2	T48154	pyruvate decarboxy
403	26	59.1	621	2	S26691	serine proteinase
404	26	59.1	625	2	C25977	phosphotransferase
405	26	59.1	625	2	T10661	serine/threonine-s
406	26	59.1	642	2	C86152	T713.2 protein -
407	26	59.1	647	2	T30892	hypotheical prote
408	26	59.1	662	2	D40228	neurexin II-beta p
409	26	59.1	665	2	PS0043	genome polypotein
410	26	59.1	675	2	A35743	creatine kinase (B
411	26	59.1	684	2	G84730	Mutator-like trans
412	26	59.1	716	2	H72763	hypotheical prote
413	26	59.1	727	2	B84847	probable CCGH-type
414	26	59.1	768	2	A42755	P-selectin precurs
415	26	59.1	768	2	I53821	P-selectin - rat
416	26	59.1	775	2	A48644	polypotein(C, E,
417	26	59.1	775	2	A47311	polypotein - deng
418	26	59.1	779	2	H71301	probable membrane-
419	26	59.1	787	2	G81692	inner membrane pro
420	26	59.1	787	2	E71537	probable 60kda inn
421	26	59.1	792	2	C32401	genome polypotein
422	26	59.1	792	2	B32401	genome polypotein
423	26	59.1	792	2	A32401	genome polypotein
424	26	59.1	795	2	P86529	60 kda inner membr
425	26	59.1	795	2	T20609	inner membrane pro
426	26	59.1	795	2	E81575	inner membrane pro
427	26	59.1	795	2	G72092	60 kda inner membr
428	26	59.1	812	2	T52569	squamosa-promoter
429	26	59.1	826	2	A60385	monocyte surfacer
430	26	59.1	858	2	JC7683	taase receptor TIR
431	26	59.1	879	2	C90879	hypotheical prote
432	26	59.1	879	2	G85739	hypotheical prote
433	26	59.1	883	2	A96662	hypotheical prote
434	26	59.1	884	2	E75489	conserved hypotet
435	26	59.1	940	2	S58135	hyphally regulated
436	26	59.1	940	2	A40985	prolectin - fruit
437	26	59.1	963	2	T26022	hypotheical prote
438	26	59.1	987	2	A54092	protein-tyrosine k
439	26	59.1	994	2	A70776	probable glnf - My
440	26	59.1	998	1	Q0BB81	BI protein - black
441	26	59.1	998	1	S41397	protein A - flock
442	26	59.1	1024	2	G83022	probable two-compo
443	26	59.1	1024	2	T27631	hypotheical prote
444	26	59.1	1030	2	H88659	protein ZC518.2 (I
445	26	59.1	1035	2	G86342	hypotheical prote
446	26	59.1	1051	2	D82428	chitodextrinase VC
447	26	59.1	1097	1	RNBY13	DNA-directed DNA p
448	26	59.1	1111	1	T26972	hypotheical prote
449	26	59.1	1127	1	GNWVD2	genome polypotein
450	26	59.1	1200	2	T43148	probable protein-t
451	26	59.1	1201	2	H86434	protein P17F8.21 (
452	26	59.1	1226	1	GNWVWP	genome polypotein
453	26	59.1	1240	1	JC5209	insulin receptor s
454	26	59.1	1242	2	JS0670	insulin receptor s
455	26	59.1	1290	2	A57190	ebnerin precursor
456	26	59.1	1375	2	T18961	FAB1 protein homol
457	26	59.1	1428	2	T08852	lustrin A - Califo
458	26	59.1	1466	2	A48216	neurexin III-alpha
459	26	59.1	1471	2	B48218	ATP-binding caaset
460	26	59.1	1474	2	D88550	neurexin III-alpha
461	26	59.1	1530	1	S01393	DNA-directed RNA p
462	26	59.1	1567	2	S11672	ice nucleation pro
463	26	59.1	1574	2	T13954	MEGF6 protein - ra
464	26	59.1	1574	2	T148216	neurexin III-alpha
465	26	59.1	1592	2	S48933	probable transport
466	26	59.1	1592	2	S48933	ras-responsive ele
467	26	59.1	1615	2	JC6510	
468	26	59.1	1630	2	A53577	aeicles sialoglyco
469	26	59.1	1639	1	NMFPB2	laminin gamma-1 ch
470	26	59.1	1650	2	S28721	hypotheical prote
471	26	59.1	1715	2	C40228	neurexin II-alpha
472	26	59.1	1732	2	T30836	lyline-specific cy
473	26	59.1	1742	2	S24600	prolectin - fruit
474	26	59.1	1766	1	NMHBUB1	laminin beta-1 cha
475	26	59.1	1786	1	NMMSB1	laminin beta-1 cha
476	26	59.1	1790	1	NMFPB1	laminin beta-1 cha
477	26	59.1	1798	1	S53869	laminin beta-2 cha
478	26	59.1	1801	1	NMRTS	laminin beta-2 cha
479	26	59.1	1816	1	S68960	laminin alpha-4 ch
480	26	59.1	1892	2	T18314	hypotheical prote
481	26	59.1	1906	1	S68235	myosin-light-chain
482	26	59.1	1955	1	AGCH	agrin precursor -
483	26	59.1	1959	1	AGRT	agrin - rat
484	26	59.1	2019	1	UQ1322	tenascin precursor
485	26	59.1	2083	2	T42721	CRP-ductin-alpha p
486	26	59.1	2090	2	T30075	hypotheical prote
487	26	59.1	2101	2	S57245	insulin receptor (
488	26	59.1	2145	2	S61041	glutamate synthase
489	26	59.1	2153	2	T30074	hypotheical prote
490	26	59.1	2344	2	T41590	probable sensor-i
491	26	59.1	2684	2	A96521	protein F2ID18.22
492	26	59.1	2844	2	S28291	hypotheical prote
493	26	59.1	3163	2	AB0233	yersinabactin bio
494	26	59.1	3163	2	T17440	probable polyketid
495	26	59.1	3388	1	GNWVDP	genome polypotein
496	26	59.1	3390	1	GNWVD3	genome polypotein
497	26	59.1	3391	1	GNWV16	genome polypotein
498	26	59.1	3391	1	GNWV26	genome polypotein
499	26	59.1	3391	1	GNWV7A	polypotein - deng
500	26	59.1	3391	2	JS0219	genome polypotein
501	26	59.1	3396	1	A42551	genome polypotein
502	26	59.1	3330	1	GNWVW	genome polypotein
503	26	59.1	3712	1	S18253	laminin alpha-1 ch
504	26	59.1	3898	1	A44217	genome polypotein
505	26	59.1	3988	1	GNWVBV	genome polypotein
506	26	59.1	5376	2	T42215	zonadhesin - mouse
507	26	59.1	6558	2	T13931	prolectin - fruit
508	26	59.1	734	2	C87400	Fold bifunctional
509	25	56.8	82	1	T31219	hypotheical prote
510	25	56.8	82	1	B46308	E3 class 2 protein
511	25	56.8	88	1	TZHQK	killer toxin - yea
512	25	56.8	89	2	B95913	hypotheical expor
513	25	56.8	100	2	A46308	E3 class 1 protein
514	25	56.8	105	2	E41377	hypotheical prote
515	25	56.8	118	1	PSL74E	phospholipase A2 (
516	25	56.8	118	1	T17967	ERV1 protein homol
517	25	56.8	130	1	KRSH3A	keratin high-sulfu
518	25	56.8	131	1	KRGT3M	keratin high-sulfu
519	25	56.8	131	1	KRSHA3	keratin high-sulfu
520	25	56.8	132	1	KRGT3J	keratin high-sulfu
521	25	56.8	132	1	S24302	hypotheical prote
522	25	56.8	132	2	S00385	hypotheical prote
523	25	56.8	133	2	I52649	circadian protein
524	25	56.8	144	2	A42585	trypsin inhibitor
525	25	56.8	145	2	D75603	hypotheical prote
526	25	56.8	145	2	UC4273	pleiotrophic facto
527	25	56.8	161	2	UC4274	pleiotrophic facto
528	25	56.8	161	2	C82703	hypotheical prote
529	25	56.8	162	2	E71368	probable RNA polym
530	25	56.8	174	2	C46308	E3 class 3 protein
531	25	56.8	183	1	I64006	hypotheical prote
532	25	56.8	188	1	UC6547	high sulfur protei
533	25	56.8	195	2	JH0719	omega-conotoxin re
534	25	56.8	202	2	E82396	malate synthase-re
535	25	56.8	204	2	T39940	lysosome (EC 3.2.1
536	25	56.8	211	2	S18463	hypotheical prote
537	25	56.8	215	2	H90856	butyrate-acetacet
538	25	56.8	218	2	B49346	ribonuclease (limp
539	25	56.8	222	2	H86394	ribonuclease (EC 3
540	25	56.8	227	2	T04420	



541	25	56.8	229	2	JC5518	614	25	56.8	421	2	T00598	hypothetical prote
542	25	56.8	230	2	D81223	615	25	56.8	421	2	JC7367	second peroxisomal
543	25	56.8	230	2	C81994	616	25	56.8	425	1	A26431	nerve growth facto
544	25	56.8	235	2	B82173	617	25	56.8	425	1	SAVYTL	satellite RNA-enco
545	25	56.8	239	2	S66341	618	25	56.8	426	2	DT1367	conserved hypothec
546	25	56.8	240	2	T04419	619	25	56.8	427	1	GQHUN	nerve growth facto
547	25	56.8	243	2	B72757	620	25	56.8	428	2	AF0302	probable solute-bi
548	25	56.8	248	2	T22755	621	25	56.8	428	2	S03767	cellulase (EC 3.2.
549	25	56.8	261	2	H83848	622	25	56.8	433	2	F81940	probable adenosylm
550	25	56.8	266	2	A12762	623	25	56.8	433	2	D81164	adenosylmethionine
551	25	56.8	271	2	H97543	624	25	56.8	440	2	A84387	oligopeptidase ABC
552	25	56.8	274	2	AC3295	625	25	56.8	443	2	P89426	protein M162.5 [im
553	25	56.8	283	2	T02287	626	25	56.8	447	2	A47430	gastrin/cholecysto
554	25	56.8	286	2	S44561	627	25	56.8	451	2	A45643	tubulin alpha chain
555	25	56.8	291	2	A13419	628	25	56.8	453	1	NMIV3	exo-alpha-sialidas
556	25	56.8	298	2	F69515	629	25	56.8	453	2	T15374	hypothetical prote
557	25	56.8	299	1	ZZZRNK	630	25	56.8	454	1	NMIV	exo-alpha-sialidas
558	25	56.8	299	2	G95932	631	25	56.8	455	2	T39537	tubulin alpha-1 ch
559	25	56.8	299	2	D95320	632	25	56.8	455	2	A25072	hypothetical prote
560	25	56.8	300	2	A81418	633	25	56.8	456	2	A85139	hypothetical prote
561	25	56.8	302	2	C64103	634	25	56.8	461	2	T02706	protein p17p6.19 f
562	25	56.8	302	2	T37514	635	25	56.8	461	2	A86435	hypothetical prote
563	25	56.8	303	1	SYRGA	636	25	56.8	462	2	S67569	hypothetical prote
564	25	56.8	303	2	C91184	637	25	56.8	463	1	S77180	glutamyL-tRNA redu
565	25	56.8	303	2	H86030	638	25	56.8	463	2	S66504	dipeptidyl-peptida
566	25	56.8	303	2	A71557	639	25	56.8	466	2	S36794	beta-1-adrenergic
567	25	56.8	304	1	S74557	640	25	56.8	468	2	B96514	hypothetical prote
568	25	56.8	304	2	A10494	641	25	56.8	469	2	D86144	protein probable u
569	25	56.8	304	2	A45989	642	25	56.8	470	1	NMIV07	exo-alpha-sialidas
570	25	56.8	304	2	T24703	643	25	56.8	470	1	NMIVXL	exo-alpha-sialidas
571	25	56.8	305	2	E83091	644	25	56.8	478	2	S04801	exo-alpha-sialidas
572	25	56.8	305	2	A41812	645	25	56.8	478	2	SS1457	probable membrane
573	25	56.8	306	2	AB1911	646	25	56.8	479	1	JN0715	3-phytase (EC 3.1.
574	25	56.8	306	2	S32834	647	25	56.8	481	2	G86144	hypothetical prote
575	25	56.8	309	2	B87576	648	25	56.8	481	2	T35149	two-component syst
576	25	56.8	311	2	B87433	649	25	56.8	484	1	A46217	GPI-anchor biosynt
577	25	56.8	314	2	I37383	650	25	56.8	485	1	A55731	gene pIG-A protein
578	25	56.8	315	2	E83645	651	25	56.8	485	2	I52484	IMP dehydrogenase
579	25	56.8	315	2	H82062	652	25	56.8	487	2	AG0349	IMP dehydrogenase
580	25	56.8	315	2	S19737	653	25	56.8	487	2	T23776	hypothetical prote
581	25	56.8	317	2	F82672	654	25	56.8	488	1	DEBCTP	IMP dehydrogenase
582	25	56.8	317	2	A12676	655	25	56.8	488	1	H64055	IMP dehydrogenase
583	25	56.8	324	2	E44221	656	25	56.8	488	2	G85894	IMP dehydrogenase
584	25	56.8	327	2	D72503	657	25	56.8	488	2	B91050	IMP dehydrogenase
585	25	56.8	327	2	G95355	658	25	56.8	490	2	AE0820	IMP dehydrogenase
586	25	56.8	330	2	G82373	659	25	56.8	495	2	B71360	hypothetical prote
587	25	56.8	331	2	AB3471	660	25	56.8	499	1	CPBHS	carboxypeptidase C
588	25	56.8	334	2	A10795	661	25	56.8	501	2	E70596	hypothetical prote
589	25	56.8	335	2	A40036	662	25	56.8	503	2	D70930	hypothetical prote
590	25	56.8	339	2	E83209	663	25	56.8	509	2	T31136	multidrug-efflux t
591	25	56.8	341	2	G97458	664	25	56.8	509	2	A86578	hypothetical prote
592	25	56.8	342	2	F70311	665	25	56.8	510	1	S43516	carboxypeptidase C
593	25	56.8	345	2	A05279	666	25	56.8	518	2	G86454	CDS protein FYJ11.
594	25	56.8	348	2	A34705	667	25	56.8	526	2	AB3433	C-di-GMP phosphodi
595	25	56.8	363	2	E97223	668	25	56.8	527	2	T41856	AcNMPV orf119 - Bo
596	25	56.8	364	2	H71440	669	25	56.8	529	2	T10388	hypothetical prote
597	25	56.8	365	2	T06693	670	25	56.8	530	2	H72864	AcOrf-119 protein
598	25	56.8	365	2	T24955	671	25	56.8	531	2	S20900	titin - mouse (fra
599	25	56.8	370	2	C96979	672	25	56.8	531	2	B84442	hypothetical prote
600	25	56.8	375	2	JX0131	673	25	56.8	532	2	C96514	hypothetical prote
601	25	56.8	385	2	T22822	674	25	56.8	544	2	T16374	membrane transport
602	25	56.8	388	2	AB3913	675	25	56.8	547	2	T41613	hypothetical prote
603	25	56.8	394	2	S58301	676	25	56.8	548	2	T22137	hypothetical prote
604	25	56.8	395	2	T00516	677	25	56.8	550	2	T26552	hypothetical prote
605	25	56.8	397	2	S74344	678	25	56.8	553	2	T14354	probable somatic e
606	25	56.8	400	2	T46383	679	25	56.8	555	2	AE2208	hypothetical prote
607	25	56.8	403	2	T06762	680	25	56.8	557	2	AD0877	probable hemolysin
608	25	56.8	404	2	S64480	681	25	56.8	578	2	A70877	probable acyl-coAS
609	25	56.8	407	2	I48739	682	25	56.8	583	1	KJHUC	steryl-sulfatase (
610	25	56.8	407	2	A84538	683	25	56.8	584	2	JC7809	sulfakinin recepto
611	25	56.8	415	2	T49840	684	25	56.8	585	2	S43572	COSB5.5 protein (c
612	25	56.8	416	1	UN0006	685	25	56.8	585	2	B88571	protein COSB5.5 [i
613	25	56.8	418	1	S28372	686	25	56.8	586	2	A53758	protein kinase C (

687	25	56.8	587	2	A49509	protein kinase C (	760	25	56.8	6805	2	S20901	clitin - rabbit (fr
688	25	56.8	591	2	S50972	Rib2 protein - yea	761	25	56.8	26926	1	I38344	clitin, cardiac mus
689	25	56.8	592	1	A30314	protein kinase C (	762	24.5	55.7	72	1	CHBC4	heat-stable entero
690	25	56.8	592	1	UC1480	protein kinase C (	763	24.5	55.7	72	1	CHBC1B	heat-stable entero
691	25	56.8	592	2	JN0877	protein kinase C (	764	24.5	55.7	293	3	S31763	chitinase (EC 3.2.
692	25	56.8	594	2	G84441	probable sucrose-p	765	24.5	55.7	301	4	B40201	artifact-warning s
693	25	56.8	645	2	T05251	probable disease r	766	24.5	55.7	330	2	T36695	probable integral
694	25	56.8	647	2	T23814	hypothetical prote	767	24.5	55.7	600	2	D84258	alcohol dehydrogen
695	25	56.8	651	2	UC7705	death receptor-6 -	768	24.5	55.7	802	2	T24293	hypothetical prote
696	25	56.8	668	2	T05257	probable disease r	769	24.5	55.7	901	2	T04108	receptor kinase ho
697	25	56.8	670	2	F88297	protein M28.1 [imp	770	24.5	55.7	949	2	T24294	hypothetical prote
698	25	56.8	679	2	AD0507	probable glycosyl	771	24.5	55.7	1042	2	A57534	mucin 5AC (clone L
699	25	56.8	704	1	DJBPT7	DNA-directed DNA p	772	24.5	55.7	1056	2	A53767	mucin MUC5B, trach
700	25	56.8	704	2	S07512	DNA-directed DNA p	773	24	54.5	61	2	S50180	pseudo-hevein - pa
701	25	56.8	709	2	S33907	sugar hydrolase (E	774	24	54.5	66	2	T17015	metallothionein-1
702	25	56.8	717	2	E86812	hypothetical prote	775	24	54.5	66	2	T17015	metallothionein-1
703	25	56.8	732	2	B69749	hypothetical prote	776	24	54.5	71	2	S39422	metallothionein 20
704	25	56.8	732	2	T05250	probable disease r	777	24	54.5	71	2	S39421	metallothionein 20
705	25	56.8	769	2	A71403	hypothetical prote	778	24	54.5	71	2	S39420	metallothionein 20
706	25	56.8	771	2	AF2381	hypothetical prote	779	24	54.5	71	2	S39419	metallothionein 20
707	25	56.8	781	2	T20864	hypothetical prote	780	24	54.5	72	2	S39418	metallothionein 10
708	25	56.8	835	2	T05259	probable disease r	781	24	54.5	72	2	S39417	metallothionein 10
709	25	56.8	844	2	B64678	NADH2 dehydrogenas	782	24	54.5	72	2	S39416	metallothionein 10
710	25	56.8	846	2	H70599	hypothetical prote	783	24	54.5	72	2	S39415	metallothionein 10
711	25	56.8	849	2	H71838	NADH2 dehydrogenas	784	24	54.5	72	2	S39414	metallothionein 10
712	25	56.8	852	2	C98310	haer protein (Y089	785	24	54.5	81	2	S27366	metallothionein B (
713	25	56.8	868	2	D86349	hypothetical prote	786	24	54.5	81	2	S27369	metallothionein B (
714	25	56.8	869	2	A71400	probable disease r	787	24	54.5	95	2	S76543	keratin, feather -
715	25	56.8	910	2	T29935	hypothetical prote	788	24	54.5	102	1	KREUB	metallothionein-1i
716	25	56.8	919	2	T32541	unc-5 protein - Ca	789	24	54.5	115	2	S16534	conserved hypotet
717	25	56.8	923	2	AH2972	unc-5 protein - Ca	790	24	54.5	120	2	F82185	high sulfur protei
718	25	56.8	947	1	B44294	unc-5 protein, lon	791	24	54.5	122	2	JC6548	hypothetical prote
719	25	56.8	1002	2	T43236	carboxypeptidase C	792	24	54.5	134	2	T22275	thi protein - Rhiz
720	25	56.8	1018	1	GNMXG7	genome polypeptid	793	24	54.5	168	2	S78110	hypothetical lipop
721	25	56.8	1170	2	A53612	laminin Blk chain	794	24	54.5	174	2	B90657	hypothetical prote
722	25	56.8	1192	2	S69000	laminin gamma 2 ch	795	24	54.5	174	2	B85508	hypothetical prote
723	25	56.8	1193	2	A44018	laminin B2t chain	796	24	54.5	177	2	H97205	hypothetical prote
724	25	56.8	1231	2	T35227	probable nitrate r	797	24	54.5	183	2	T49855	hypothetical prote
725	25	56.8	1245	2	T42820	hypothetical prote	798	24	54.5	195	2	T28803	hypothetical prote
726	25	56.8	1246	2	C36806	hypothetical prote	799	24	54.5	200	2	F69047	hypothetical prote
727	25	56.8	1302	2	T00038	probable RNA-direc	800	24	54.5	201	2	D71190	hypothetical prote
728	25	56.8	1309	2	T00078	hypothetical prote	801	24	54.5	202	2	T32159	hypothetical prote
729	25	56.8	1321	2	T00382	hypothetical prote	802	24	54.5	235	2	D65071	hypothetical prote
730	25	56.8	1376	2	G00043	osteonidogen - hum	803	24	54.5	235	2	A85943	hypothetical prote
731	25	56.8	1378	2	T47505	RING finger-like p	804	24	54.5	235	2	E91097	hypothetical prote
732	25	56.8	1490	2	UC5145	DNA (cytosine-5-) -	805	24	54.5	241	1	CCBO1	ubiquinol-cytochro
733	25	56.8	1495	2	S22610	DNA (cytosine-5-) -	806	24	54.5	245	2	F90823	probable tail asse
734	25	56.8	1502	1	RGBYH1	CYC1/CYP3 transcri	807	24	54.5	245	2	C85682	probable tail asse
735	25	56.8	1537	2	JC4172	DNA (cytosine-5-) -	808	24	54.5	247	2	B90877	probable tail asse
736	25	56.8	1550	2	S60328	glutamate synthase	809	24	54.5	247	2	F90897	tail assembly prot
737	25	56.8	1607	1	MMMSB2	laminin gamma-1 ch	810	24	54.5	247	2	F90834	hypothetical prote
738	25	56.8	1609	1	MMHUB2	laminin gamma-1 ch	811	24	54.5	270	2	T47421	hypothetical prote
739	25	56.8	1612	2	UC5210	DNA (cytosine-5-) -	812	24	54.5	272	2	H83269	hypothetical prote
740	25	56.8	1622	2	JEO378	adhesive ligand ep	813	24	54.5	272	2	T24709	hypothetical prote
741	25	56.8	1713	2	A55347	MEGF3 protein - hu	814	24	54.5	275	2	B55224	hypothetical prote
742	25	56.8	1737	2	T00209	laminin alpha-2 ch	815	24	54.5	280	2	T09504	LIM-protein FHL3,
743	25	56.8	1751	1	MMHUMH	hypothetical prote	816	24	54.5	282	1	YPDOD1	prestalk Dll prote
744	25	56.8	1808	2	T15099	hypothetical prote	817	24	54.5	289	2	F71402	xyloglucan endo-1,
745	25	56.8	1945	2	T13937	Plexin A - fruit f	818	24	54.5	289	2	S04648	ampr protein - Rho
746	25	56.8	2116	2	T49818	glutamate synthase	819	24	54.5	289	2	T49024	hypothetical prote
747	25	56.8	2150	2	T32497	hypothetical prote	820	24	54.5	293	2	A85817	hypothetical prote
748	25	56.8	2356	2	T27790	hypothetical prote	821	24	54.5	293	2	E85718	probable tail comp
749	25	56.8	2535	2	T02546	hypothetical prote	822	24	54.5	300	2	T49748	hypothetical prote
750	25	56.8	2823	2	F87908	protein T22A3.8 [i	823	24	54.5	302	2	AB0857	ATP sulfurylase (A
751	25	56.8	2823	2	T23064	hypothetical prote	824	24	54.5	302	2	D65056	ATP sulfurylase su
752	25	56.8	3075	2	S14458	laminin alpha-1 ch	825	24	54.5	302	2	G85924	ATP sulfurylase su
753	25	56.8	3084	1	MMMSA	laminin alpha-1 ch	826	24	54.5	302	2	F91079	sulfate adenylylctr
754	25	56.8	3102	2	T43391	laminin alpha-1 ch	827	24	54.5	302	2	AH0408	hypothetical prote
755	25	56.8	3106	1	S53868	laminin alpha-2 ch	828	24	54.5	302	2	T32872	hypothetical prote
756	25	56.8	3672	2	T23433	hypothetical prote	829	24	54.5	303	2	A86443	probable major int
757	25	56.8	3704	2	T37316	probable laminin a	830	24	54.5	313	2	A11948	hypothetical prote
758	25	56.8	3707	2	S18252	heparan sulfatase pr	831	24	54.5	320	2	A53119	cell adhesion glyco
759	25	56.8	4391	2	A38096	perlecan precursor	832	24	54.5	325	1	S00680	ubiquinol-cytochro

833	24	54.5	311	2	T04454	probable peroxidase	906	23.5	53.4	492	2	A61382	phosphorylation re
834	24	54.5	317	2	T41962	hypothetical prote	907	23.5	53.4	527	2	C64139	translation releas
835	24	54.5	355	2	S74204	C4-dicarboxylate m	908	23.5	53.4	529	2	AD1072	peptide chain rele
836	24	54.5	356	2	T04360	probable peroxidase	909	23.5	53.4	529	2	E91295	peptide chain rele
837	24	54.5	353	2	C64098	tRNA (uracil-5-)-m	910	23.5	53.4	529	2	G66136	peptide chain rele
838	24	54.5	366	2	A37321	tRNA (uracil-5-)-m	911	23.5	53.4	529	2	AE0053	peptide chain rele
839	24	54.5	366	2	H91240	tRNA (uracil-5-)-m	912	23.5	53.4	529	2	I59305	translation releas
840	24	54.5	366	2	AC0935	tRNA (uracil-5-)-me	913	23.5	53.4	531	2	E82295	translation releas
841	24	54.5	366	2	EB6088	tRNA (uracil-5-)-m	914	23.5	53.4	27	2	T12330	metallochionein -
842	24	54.5	367	2	AE0476	tRNA (uracil-5-)-m	915	23.5	53.4	46	1	VTVA03	viscotoxin A3 - Bu
843	24	54.5	373	2	T29596	hypothetical prote	916	23.5	52.3	46	1	DKOGB	viscotoxin A3 - Bu
844	24	54.5	380	2	D44490	retrovirus-related	917	23.5	52.3	46	1	LORDAL	ligetoxin A - Arge
845	24	54.5	384	2	D69809	spore germination	918	23.5	52.3	46	1	VFPDIT	phoratoxin A - Bue
846	24	54.5	392	2	B48423	homeotic protein e	919	23.5	52.3	46	1	VTVA02	viscotoxin A2 - Bu
847	24	54.5	401	2	A48423	engrailed homeodom	920	23.5	52.3	46	1	VTVA02	viscotoxin B - Eur
848	24	54.5	407	2	T41708	gtp binding prote	921	23.5	52.3	52	2	D95920	hypothetical prote
849	24	54.5	427	2	S71367	u1 snRNP 70K prote	922	23.5	52.3	57	1	TIRPVB	venom basic protei
850	24	54.5	435	2	T08038	probable mitochond	923	23.5	52.3	59	2	I51272	retinoic acid rece
851	24	54.5	450	2	G01158	tyrosine kinase ac	924	23.5	52.3	63	2	S07127	chymotrypsin/elast
852	24	54.5	468	2	T03164	probable major env	925	23.5	52.3	65	2	T09787	metallochionein-11
853	24	54.5	484	2	AC3022	glutamate synthase	926	23.5	52.3	65	4	C40499	defensin alpha-3 p
854	24	54.5	484	2	G98262	glutamate synthase	927	23.5	52.3	67	1	TIROC	trypsin inhibitor,
855	24	54.5	499	1	P1ME33	li protein - human	928	23.5	52.3	75	1	GSRF8	salivary glue prot
856	24	54.5	499	2	AE3514	glutamate synthase	929	23.5	52.3	80	2	AE2958	hypothetical prote
857	24	54.5	506	1	CSFP	catalase (EC 1.11.	930	23.5	52.3	80	2	P96324	hypothetical prote
858	24	54.5	507	2	S64507	probable membrane	931	23.5	52.3	97	2	T38323	very hypothetical
859	24	54.5	516	2	T35542	probable large sec	932	23.5	52.3	97	2	B25439	Beelin protein - s
860	24	54.5	521	2	H69317	conserved hypotet	933	23.5	52.3	105	2	F72614	hypothetical prote
861	24	54.5	524	1	P1ME58	li protein - human	934	23.5	52.3	111	2	S16099	viscotoxin - Europ
862	24	54.5	532	1	C97228	probable peptide A	935	23.5	52.3	111	2	E72648	hypothetical prote
863	24	54.5	538	2	G69317	conserved hypotet	936	23.5	52.3	122	2	Q00150	hypothetical 13K p
864	24	54.5	565	2	T14732	probable beta-gluc	937	23.5	52.3	133	2	S52547	thionin variant Th
865	24	54.5	602	1	S25116	hydroxymethylgluta	938	23.5	52.3	133	2	S52554	thionin variant Th
866	24	54.5	618	1	B43255	hydrogenase (EC 1.	939	23.5	52.3	134	2	S52554	thionin variant Th
867	24	54.5	618	2	S09251	hydrogenase (EC 1.	940	23.5	52.3	134	2	T49623	hypothetical prote
868	24	54.5	619	1	JH0776	hydrogenase (EC 1.	941	23.5	52.3	135	2	H84523	probable thionin l
869	24	54.5	622	1	OYCHGC	phosphoenolpyruvat	942	23.5	52.3	136	2	I56339	cytochrome-c oxida
870	24	54.5	622	1	OYRGP	phosphoenolpyruvat	943	23.5	52.3	141	2	S73027	hypothetical prote
871	24	54.5	622	1	A45746	phosphoenolpyruvat	944	23.5	52.3	148	2	F71454	hypothetical prote
872	24	54.5	680	2	A28121	major merocrite su	945	23.5	52.3	152	1	KRSHHC	keratin high-sulfu
873	24	54.5	820	2	C81252	probable NMDH2 deh	946	23.5	52.3	152	2	I47111	high-sulfur wool m
874	24	54.5	829	2	A40894	RNA-directed RNA p	947	23.5	52.3	152	2	I47109	high-sulfur wool m
875	24	54.5	886	2	A57172	probable hormone r	948	23.5	52.3	152	2	I47108	high-sulfur wool m
876	24	54.5	912	2	JH0563	metabotropic gluta	949	23.5	52.3	152	2	I47112	high-sulfur wool m
877	24	54.5	956	1	A46016	thrombospondin 3 -	950	23.5	52.3	156	1	KRSHHB	keratin high-sulfu
878	24	54.5	956	2	A57121	thrombospondin 3 p	951	23.5	52.3	157	2	JN0057	hypothetical 17K p
879	24	54.5	957	2	T15976	hypothetical prote	952	23.5	52.3	158	2	T12958	copper transport p
880	24	54.5	1036	2	T38734	hypothetical prote	953	23.5	52.3	160	2	A69849	hypothetical prote
881	24	54.5	1043	2	T19734	hypothetical prote	954	23.5	52.3	162	2	I47107	high-sulfur wool m
882	24	54.5	1166	1	S06142	protein-tyrosine k	955	23.5	52.3	162	2	C81808	hypothetical prote
883	24	54.5	1173	1	VG1HHC	E2 glycoprotein pr	956	23.5	52.3	163	2	H97034	hypothetical secre
884	24	54.5	1225	1	S24284	E2 glycoprotein pr	957	23.5	52.3	165	2	B71432	hypothetical prote
885	24	54.5	1225	1	A36607	E2 glycoprotein -	958	23.5	52.3	169	1	S18946	ultra high-sulfur
886	24	54.5	1225	2	S55598	tegment protein 0	959	23.5	52.3	172	1	KRSHHA	keratin high-sulfu
887	24	54.5	1360	2	T33922	hypothetical prote	960	23.5	52.3	172	2	I47106	high-sulfur wool m
888	24	54.5	1414	2	T33236	hypothetical prote	961	23.5	52.3	179	2	T49508	hypothetical prote
889	24	54.5	1447	1	VG1HE3	E2 glycoprotein pr	962	23.5	52.3	182	1	KRSHHD	keratin high-sulfu
890	24	54.5	1447	1	VG1HE3	E2 glycoprotein pr	963	23.5	52.3	182	2	I47105	ultra-high-sulfur
891	24	54.5	1449	1	A43573	E2 glycoprotein pr	964	23.5	52.3	182	2	A36686	hypothetical prote
892	24	54.5	1449	1	VG1HFS	E2 glycoprotein pr	965	23.5	52.3	186	2	T32656	hypothetical prote
893	24	54.5	1449	2	S47423	E2 glycoprotein pr	966	23.5	52.3	187	2	T38452	hypothetical prote
894	24	54.5	1451	1	JOL1719	E2 glycoprotein pr	967	23.5	52.3	191	2	I46412	keratin KAP5.4 - s
895	24	54.5	1453	1	S41453	spike protein - ca	968	23.5	52.3	195	2	T03522	bluf protein - Rho
896	24	54.5	1506	2	T30886	integumentary muc	969	23.5	52.3	195	2	S52221	bluf protein - Rho
897	24	54.5	1534	2	T30295	p-glycoprotein - T	970	23.5	52.3	196	2	D63275	conserved hypotet
898	24	54.5	1565	2	AD2135	polyketide synthas	971	23.5	52.3	197	2	I46413	keratin KAP5.5 - s
899	24	54.5	1587	2	G86467	hypothetical prote	972	23.5	52.3	198	2	S25656	T-cell surface gly
900	24	54.5	1721	2	T21214	hypothetical prote	973	23.5	52.3	199	2	E75463	hypothetical prote
901	24	54.5	1772	2	A45532	major merocrite su	974	23.5	52.3	205	2	T14745	hypothetical prote
902	23.5	53.4	178	2	T38624	hypothetical prote	975	23.5	52.3	208	2	T34512	hypothetical prote
903	23.5	53.4	333	2	T32058	hypothetical prote	976	23.5	52.3	219	2	T21736	hypothetical prote
904	23.5	53.4	464	2	S50592	hypothetical prote	977	23.5	52.3	235	1	RWHUT8	T-cell surface gly
905	23.5	53.4	468	2	S39832	probable phosphopr	978	23.5	52.3	238	2	I48605	insulin-like growt

979 23 52.3 241 2 T29182 hypothetical prote  
980 23 52.3 244 2 UC8019 CDS6 protein p3g  
981 23 52.3 250 1 D64740 probable phosphos  
982 23 52.3 250 1 DGBEL5 uracil-DNA glycosy  
983 23 52.3 258 2 T44251 creatinase (EC 3  
984 23 52.3 260 1 OTXL3 cytochrome-c oxida  
985 23 52.3 260 2 T11133 cytochrome-c oxida  
986 23 52.3 261 1 OTBO3 cytochrome-c oxida  
987 23 52.3 261 1 OTCA3 cytochrome-c oxida  
988 23 52.3 261 1 OTMS3 cytochrome-c oxida  
989 23 52.3 261 2 T11499 cytochrome-c oxida  
990 23 52.3 261 2 H34012 cytochrome-c oxida  
991 23 52.3 261 2 B30396 cytochrome-c oxida  
992 23 52.3 261 2 T10396 cytochrome-c oxida  
993 23 52.3 261 2 T11540 cytochrome-c oxida  
994 23 52.3 261 2 G58888 cytochrome-c oxida  
995 23 52.3 261 2 S47876 cytochrome-c oxida  
996 23 52.3 261 2 T11306 cytochrome-c oxida  
997 23 52.3 261 2 T11434 cytochrome-c oxida  
998 23 52.3 261 2 T11447 cytochrome-c oxida  
999 23 52.3 261 2 S60277 cytochrome-c oxida  
1000 23 52.3 261 2 D30396 cytochrome-c oxida

## ALIGNMENTS

RESULT 1  
A27020  
DIF-induced presteak pdd63 protein precursor - slime mold (Dictyostelium discoideum) (F  
C/Species: Dictyostelium discoideum  
C/Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 09-Jul-2004  
C/Accession: A27020  
R.Williams, J.G.; Ceccarelli, A.; McRobbie, S.; Mahbubani, H.; Kay, R.R.; Barly, A.; Ben  
Cell 49, 185-192, 1987  
A/Title: Direct induction of Dictyostelium presteak gene expression by DIF provides evid  
A/Reference number: A27020; PMID:87187613; PMID:3568124  
A/Accession: A27020  
A/Molecule type: DNA  
A/Residues: 1-601 <WIL>  
A/Cross-references: UNIPROT:Q7M4U3; UNIPARC:UPI000017B181  
C/Genetics:  
A/Genes: pdd63  
F/1-20/Domain: signal sequence #status predicted <SIG>  
F/21-601/Product: DIF-induced presteak pdd63 protein #status predicted <MAT>  
Query Match 81.8%; Score 36; DB 2; Length 601;  
Best Local Similarity 71.4%; Pred. No. 23;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CXPXTGC 10  
DB 245 CSPSTGC 251  
RESULT 2  
S07983  
phospholipase A2 homolog precursor - common tiger snake  
C/Species: Notechis scutatus scutatus (common tiger snake, mainland tiger snake)  
C/Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 05-Oct-2004  
C/Accession: S07983  
R.Ducancel, F.  
Submitted to the EMBL Data Library, January 1989  
A/Reference number: S07983  
A/Accession: S07983  
A/Molecule type: mRNA  
A/Residues: 1-145 <DUC>  
A/Cross-references: UNIPROT:P20146; UNIPARC:UPI0000131165; EMBL:X14043; NID:964109; PIDX  
C/Suprafamily: Phospholipase A2  
F/1-27/Domain: signal sequence #status predicted <SIG>  
F/28-145/Product: phospholipase A2 #status predicted <MAT>  
Query Match 77.3%; Score 34; DB 2; Length 145;

Best Local Similarity 71.4%; Pred. No. 18;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 105 CDPPTGC 111

## RESULT 3

T05348  
hypothetical protein F8B4.80 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C/Accession: T05348  
R.Bevan, M.; Terry, N.; Ardles, W.; Buysaert, C.; Dasseville, R.; De Clerck, R.; De  
ewes, H.W.; Mayer, K.F.X.; Schueller, C.  
Submitted to the Protein Sequence Database, February 1999  
A/Reference number: Z15409  
A/Accession: T05348  
A/Molecule type: DNA  
A/Residues: 1-312 <BEV>  
A/Cross-references: UNIPROT:Q9SU73; UNIPARC:UPI0000048A40; EMBL:AL034567  
A/Experimental source: cultivar Columbia; BAC clone F8B4  
C/Genetics:  
A/Map position: 4  
A/Introns: 1/3; 44/3; 101/3; 139/3; 180/3  
A/Note: F8B4.80  
Query Match 77.3%; Score 34; DB 2; Length 312;  
Best Local Similarity 71.4%; Pred. No. 32;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 208 CDPNTGC 214

## RESULT 4

T50496  
hypothetical protein T22D6.50 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 09-Jul-2004  
C/Accession: T50496  
R.Bevan, M.; Terry, N.; Ardles, W.; Buysaert, C.; Dasseville, R.; De Clerck, R.; De  
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.  
Submitted to the Protein Sequence Database, May 2000  
A/Reference number: Z25101  
A/Accession: T50496  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1058 <BEV>  
A/Cross-references: UNIPROT:Q9LEZ5; UNIPARC:UPI000009F613; EMBL:AL357612  
A/Experimental source: cultivar Columbia; BAC clone T22D6  
C/Genetics:  
A/Map position: 5  
A/Introns: 60/3; 195/3; 222/3; 448/3; 492/3; 526/3; 555/2; 591/2; 616/3; 662/2; 715/3; 7  
A/Note: T22D6.50  
Query Match 75.0%; Score 33; DB 2; Length 1058;  
Best Local Similarity 71.4%; Pred. No. 13e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 1019 CDPSTGC 1025

## RESULT 5

F96006  
hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplaemid pSymB  
C/Species: Sinorhizobium meliloti  
C/Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C/Accession: F96006

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernat  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo  
A:Reference number: A95842; MUID:2136508; PMID:11481431  
A:Accession: F96006  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-64 <KUR>  
A:Cross-references: UNIPROT:Q92U25; UNIPARC:UPI00000CB812; GB:AL591985; PTDN:CAC49718.1;  
R:Gallbert, F.; Finan, T.M.; Long, S.R.; Punler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: SMB21688  
A:Genome: plasmid

Query Match 72.7%; Score 32; DB 2; Length 64;  
Best Local Similarity 57.1%; Pred. No. 22;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTCG 10  
DB 15 CAPTGGC 21

## RESULT 6

T48397  
S-Receptor kinase-like protein - *Arabidopsis thaliana*  
N:Alternate names: protein F17C15.120  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 31-Dec-2004  
C:Accession: T48397  
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24492  
A:Accession: T48397  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-482 <BEV>  
A:Cross-references: UNIPROT:Q9LZR8; UNIPARC:UPI000048ABC; EMBL:AL162506  
A:Experimental source: cultivar Columbia; BAC clone F17C15  
C:Genetics:  
A:Map position: 5  
A:Note: F17C15.120  
C:Superfamily: S-locus glycoprotein

Query Match 70.5%; Score 31; DB 2; Length 482;  
Best Local Similarity 57.1%; Pred. No. 1.6e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTCG 10  
DB 301 CTPGSGC 307

## RESULT 7

T25473  
Hypothetical protein B0507.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T25473  
R:Bradsaw, H.  
submitted to the EMBL Data Library, July 1996  
A:Description: The sequence of C. elegans cosmid B0507.  
A:Reference number: Z20039  
A:Accession: T25473

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-643 <BRA>  
A:Cross-references: UNIPROT:Q22886; UNIPARC:UPI000017B6F9; EMBL:U64833; PTDN:AB04822.1.  
A:Experimental source: strain Bristol N2; clone B0507  
C:Genetics:  
A:Gene: CSPP:B0507.1  
A:Map position: 5  
A:Introns: 59/3; 133/1; 464/3; 586/1

Query Match 70.5%; Score 31; DB 2; Length 643;  
Best Local Similarity 57.1%; Pred. No. 2e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTCG 10  
DB 209 CSPQSGC 215

## RESULT 8

A24255  
Chorion class A protein L11 precursor - *silkworm*  
C:Species: *Bombyx mori* (silkworm)  
C>Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
C:Accession: A24255  
R:Spoerel, N.; Nguyen, H.T.; Kafatos, F.C.  
J. Mol. Biol. 190, 23-35, 1986  
A:Title: Gene regulation and evolution in the chorion locus of *Bombyx mori*. Structural  
A:Reference number: A92929; MUID:87060979; PMID:3023635  
A:Accession: A24255  
A:Molecule type: DNA  
A:Residues: 1-129 <SPO>  
A:Cross-references: UNIPROT:P08826; UNIPARC:UPI0000127760; GB:X15557; GB:X04028; GB:X040  
C:Superfamily: chorion class A protein pc292  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-129/Product: chorion class A protein L11 #status predicted <MAT>

Query Match 68.2%; Score 30; DB 2; Length 129;  
Best Local Similarity 57.1%; Pred. No. 88;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTCG 10  
DB 116 CAPTCGC 122

## RESULT 9

B24255  
Chorion class A protein L12 precursor - *silkworm*  
C:Species: *Bombyx mori* (silkworm)  
C>Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
C:Accession: B24255  
R:Spoerel, N.; Nguyen, H.T.; Kafatos, F.C.  
J. Mol. Biol. 190, 23-35, 1986  
A:Title: Gene regulation and evolution in the chorion locus of *Bombyx mori*. Structural  
A:Reference number: A92929; MUID:87060979; PMID:3023635  
A:Accession: B24255  
A:Molecule type: DNA  
A:Residues: 1-132 <SPO>  
A:Cross-references: UNIPROT:P08825; UNIPARC:UPI0000127761; GB:X15557; GB:X04028; GB:X040  
C:Superfamily: chorion class A protein pc292  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-132/Product: chorion class A protein L12 #status predicted <MAT>

Query Match 68.2%; Score 30; DB 2; Length 132;  
Best Local Similarity 57.1%; Pred. No. 89;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTCG 10  
DB 117 CAPTCGC 123

RESULT 10  
AB0331  
A:Sigma E factor regulatory protein rsec [imported] - Yersinia pestis (strain C092)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AB0331  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AB0331  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-151 <KUR>  
A:Cross-references: UNIPROT:Q8ZD76; UNIPARC:UP100000DC88E; GB:AL590842; PIDN:CMC92953.1;  
C:Genetics:  
A:Gene: rsec  
C:Superfamily: Escherichia coli sigma-E factor regulatory protein rsec  
Query Match 68.2%; Score 30; DB 2; Length 151;  
Best Local Similarity 57.1%; Pred. No. 99;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CXPXTGC 10  
DB 20 CEPHSGC 26  
RESULT 11  
T99486  
A:hypothetical protein B14D6.380 [imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000  
C:Accession: T49486  
R:Schulte, U.; Align, V.; Hohesl, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A:Reference number: Z25022  
A:Accession: T49486  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-119 <SCH>  
A:Cross-references: UNIPARC:UP10000179E18; EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.380  
A:Experimental source: BAC clone B14D6; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B14D6.380  
A:Map position: 6  
C:Superfamily: Neurospora crassa hypothetical protein B14D6.380  
Query Match 68.2%; Score 30; DB 2; Length 199;  
Best Local Similarity 57.1%; Pred. No. 1.2e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 4 CXPXTGC 10  
DB 155 CKPLSGC 161  
RESULT 12  
S04926  
A:wound-induced protein 1 precursor - potato  
C:Species: Solanum tuberosum (potato)  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 09-Jul-2004  
C:Accession: S04926  
R:Stanford, A.; Bevan, M.; Northcote, D.  
Mol. Gen. Genet. 215, 200-208, 1989  
A:Title: Differential expression within a family of novel wound-induced genes in potato.  
A:Reference number: S04926; MUID:89218921; PMID:2710099  
A:Molecule type: DNA  
A:Residues: 1-200 <STA>  
A:Cross-references: UNIPROT:P09761; UNIPARC:UP10000138F0C; EMBL:X13497; NID:G21617; PIDN

C:Genetics:  
A:Gene: win1  
A:Accession: 143/3  
C:Superfamily: hevein precursor; barwin homology; hevein chitin-binding domain homology  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-200/Product: wound-induced protein 1 #status predicted <MAT>  
F:26-69/Domain: hevein chitin-binding domain homology <HCB>  
F:78-199/Domain: barwin homology <BAR>  
Query Match 68.2%; Score 30; DB 2; Length 200;  
Best Local Similarity 57.1%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 4 CXPXTGC 10  
DB 56 CSPSGC 62  
RESULT 13  
T07729  
A:wound-induced protein (clone TAB7) - tomato (fragment)  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004  
C:Accession: T07729  
R:Harris, N.; Taylor, J.E.; Roberts, J.A.  
J. Exp. Bot. 48, 1223-1227, 1997  
A:Title: Characterization and expression of an mRNA encoding a wound-induced (win) prote  
A:Reference number: Z16099  
A:Accession: T07729  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-201 <HAR>  
A:Cross-references: UNIPROT:003994; UNIPARC:UP100000AC92; EMBL:U89764; NID:G1888560; PI  
A:Experimental source: strain Alisa Craig; leaf abscission zone tissue  
C:Superfamily: hevein precursor; barwin homology; hevein chitin-binding domain homology  
F:16-60/Domain: hevein chitin-binding domain homology <HCB>  
F:68-188/Domain: barwin homology <BAR>  
Query Match 68.2%; Score 30; DB 2; Length 201;  
Best Local Similarity 57.1%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 4 CXPXTGC 10  
DB 46 CSPSGC 52  
RESULT 14  
T04098  
A:CBP20 preprotein - common tobacco  
N:Alternate names: wound-induced protein  
C:Species: Nicotiana tabacum (common tobacco)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T04098  
R:Ponstein, A.S.; Bres-Vioemann, S.A.; Sela-Buurlage, W.B.; Elzen, P.J.; Melchers, L.S.,  
Plant Physiol. 104, 109-118, 1994  
A:Title: A novel pathogen- and wound-inducible tobacco (Nicotiana tabacum) protein with  
A:Reference number: Z15209; MUID:94159785; PMID:8115541  
A:Accession: T04098  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-211 <PON>  
A:Cross-references: UNIPROT:Q41231; UNIPARC:UP100000A8993; EMBL:S72452; NID:G632733; PID  
A:Experimental source: cultivar Samum NN  
C:Superfamily: hevein precursor; barwin homology; hevein chitin-binding domain homology  
F:23-66/Domain: hevein chitin-binding domain homology <HCB>  
F:78-199/Domain: barwin homology <BAR>  
Query Match 68.2%; Score 30; DB 2; Length 211;  
Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 4 CXPXTGC 10

Db 53 CSDPGC 59

## RESULT 15

S04927  
C:Species: Solanum tuberosum (potato)  
C>Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 09-Jul-2004  
C/Accession: S04927  
R/Stanford, A.; Bevan, M.; Northcote, D.  
Mol. Gen. Genet. 215, 200-208, 1989  
A>Title: Differential expression within a family of novel wound-induced genes in potato.  
A/Reference number: S04926; PMID:89218921; PMID:2710099  
A/Accession: S04927  
A:Molecule type: DNA  
A:Residues: 1-211 <STA>  
A/Cross-references: UNIPROT:P09762; UNIPARC:UPI0000138F0D; EMBL:X13497; NID:g21617; PIDD:  
C/Genetics:  
A:Gene: win2  
A:Introns: 142/3  
C:Superfamily: hevein precursor; barwin homology; hevein chitin-binding domain homology  
F:/25/Domain: signal sequence #status predicted <SIG>  
F:/26-211/Product: wound-induced protein 2 #status predicted <MAT>  
F:/26-69/Domain: hevein chitin-binding domain homology <HCB>  
F:/77-198/Domain: barwin homology <BAR>

## Query Match

Best Local Similarity 68.2%; Score 30; DB 2; Length 211;  
Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
Db 56 CSDPGC 62

## RESULT 16

MMVZB4  
C:Species: vaccinia virus (strain Copenhagen)  
N/Alternate names: B14R protein; serine proteinase inhibitor  
A/Note: host Homo sapiens (man)  
C/Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 05-Oct-2004  
C/Accession: D42527  
R/Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Winslow, J.P.; Paolletti, E.  
Virology 179, 517-563, 1990  
A>Title: Appendix to "The complete DNA sequence of vaccinia virus".  
A/Reference number: A42501  
A/Accession: D42527  
A:Molecule type: DNA  
A:Residues: 1-222 <GOE>  
A/Cross-references: UNIPROT:P20842; UNIPARC:UPI0000135E37; GB:M35027; NID:g335317; PIDD:  
A/Experimental source: strain Copenhagen  
R/Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Winslow, J.P.; Paolletti, E.  
Virology 179, 247-266, 1990  
A>Title: The complete DNA sequence of vaccinia virus.  
A/Reference number: A42531; PMID:91021027; PMID:2219722  
A/Contents: annotation; possible protein-coding frames  
A/Note: neither amino acid nor nucleotide sequence is given  
C:Superfamily: serpin  
C/Keywords: glycoprotein; serine proteinase inhibitor  
F:/62,93/Binding site: carbohydrate (Asn) (covalent) #status predicted

## Query Match

Best Local Similarity 68.2%; Score 30; DB 1; Length 222;  
Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
Db 216 CSDPGC 222

## RESULT 17

E84326  
hypothetical protein Vng1748c [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 05-Oct-2004  
C/Accession: E84326  
R/Ng, W.V.; Kennedy, S.P.; Mahatae, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laeky,  
; Leithauer, B.; Keller, K.; Cruz, R.; Danison, M.J.; Hough, D.W.; Maddocks, D.G.; Jabic  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;  
A>Title: Genome sequence of Halobacterium species NRC-1.  
A/Reference number: E84326  
A/Accession: E84326  
A/Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-224 <STO>  
A/Cross-references: UNIPROT:Q9HP93; UNIPARC:UPI0000063970; GB:AE004437; NID:g10581204; I  
C/Genetics:  
A:Gene: VNG1748C  
C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology

## Query Match

Best Local Similarity 68.2%; Score 30; DB 2; Length 224;  
Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
Db 168 CSDPGC 174

## RESULT 18

A83801  
ABC transporter (substrate-binding protein) BH1209 [imported] - Bacillus halodurans (str  
C:Species: Bacillus halodurans  
C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C/Accession: A83801  
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maeno, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A/Reference number: A83650; PMID:20512582; PMID:11058132  
A/Accession: A83801  
A/Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-336 <STO>  
A/Cross-references: UNIPROT:Q9KDK4; UNIPARC:UPI00000D73E9; GB:AP001511; GB:BA000004; NID  
A/Experimental source: strain C-125  
C/Genetics:  
A:Gene: BH1209

## Query Match

Best Local Similarity 68.2%; Score 30; DB 2; Length 336;  
Best Local Similarity 57.1%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
Db 139 CSDPGC 145

## RESULT 19

MMVZM2  
C:Species: vaccinia virus  
N/Alternate names: B13R protein; serine proteinase inhibitor 2  
A/Note: host Homo sapiens (man)  
C/Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 05-Oct-2004  
C/Accession: A44035; J01807; B30175; C42527  
R/Smith, G.L.; Howard, S.T.; Chan, Y.S.  
J. Gen. Virol. 70, 2333-2343, 1989  
A>Title: Vaccinia virus encodes a family of genes with homology to serine proteinase inh  
A/Reference number: A44035; PMID:89381686; PMID:2778436  
A/Accession: A44035  
A:Molecule type: DNA  
A:Residues: 1-345 <SMT>  
A/Cross-references: UNIPROT:P15059; UNIPARC:UPI0000135B2A; EMBL:D00581; NID:g222694; PID

A:Experimental source: strain WR  
 R:Smith, G.L.; Chan, Y.S.; Howard, S.T.  
 J:Gen. Virol. 72, 1349-1376, 1991  
 A:Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right in  
 A:Reference number: JQ1767; MUID:91259063; PMID:2045793  
 A:Accession: JQ1807  
 A:Molecule type: DNA  
 A:Residues: 1-345 <SN>  
 A:Cross-references: UNIPARC:UPI0000135E2A; DDBJ:D11079; NID:g222717; PIDN:BAA01843.1; PI  
 A:Experimental source: strain WR  
 R:Kotwal, G.J.; Moss, B.  
 J:Virol. 63, 600-606, 1989  
 A:Title: Vaccinia virus encodes two proteins that are structurally related to members of  
 A:Reference number: A30175; MUID:89094985; PMID:2783466  
 A:Accession: B30175  
 A:Molecule type: DNA  
 A:Residues: 1-56, 'V', 61-129, 'S', 131-133, 'DV', 136-164, 'CI', 167-171, 'E', 173-179, 'T', 181-19  
 A:Cross-references: UNIPARC:UPI0000173205  
 R:Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.  
 Virol. 179, 517-563, 1990  
 A:Title: Appendix to "The complete DNA sequence of vaccinia virus".  
 A:Reference number: A42501  
 A:Accession: C42527  
 A:Molecule type: DNA  
 A:Residues: 11-47, 'T', 49-51, 'T', 53-78, 'P', 80-81, 'G', 83-124, 'SN' <GOE>  
 A:Cross-references: UNIPARC:UPI0000135E35; GB:M35027; NID:g335317; PIDN:AAA48210.1; PID:  
 A:Experimental source: strain Copenhagen  
 R:Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paolletti, E.  
 Virol. 179, 247-266, 1990  
 A:Title: The complete DNA sequence of vaccinia virus.  
 A:Reference number: A42531; MUID:91021027; PMID:2219722  
 A:Contents: annotation; possible protein-coding frames  
 A:Note: neither amino acid nor nucleotide sequence is given  
 C:Genetics:  
 A:Gene: SPI-2  
 C:Superfamily: serpin  
 C:Keywords: glycoprotein; serine proteinase inhibitor  
 F:35,59,185,216/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 68.2%; Score 30; DB 1; Length 345;  
 Best Local Similarity 57.1%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXGCG 10  
 DB 339 CSPTTNC 345

RESULT 20  
 T23898  
 hypothetical protein R04D3.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T23898  
 R:Swindburne, J.  
 submitted to the EMBL Data Library, March 1996  
 A:Reference number: Z19815  
 A:Accession: T23898  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-406 <WIL>  
 A:Cross-references: UNIPROT:Q21712; UNIPARC:UPI00000766B6; EMBL:Z70212; PIDN:CAA94162.1;  
 A:Experimental source: clone R04D3  
 C:Genetics:  
 A:Gene: CSP:R04D3.2  
 A:Map position: X  
 A:Introns: 108/3; 320/3

Query Match 68.2%; Score 30; DB 2; Length 406;  
 Best Local Similarity 57.1%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXGCG 10

DB 36 CSPNAGC 42

RESULT 21  
 T23899  
 hypothetical protein R04D3.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T23899  
 R:Swindburne, J.  
 submitted to the EMBL Data Library, March 1996  
 A:Reference number: Z19815  
 A:Accession: T23899  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-410 <WIL>  
 A:Cross-references: UNIPROT:Q21713; UNIPARC:UPI000007D6C1; EMBL:Z70212; PIDN:CAA94163.1;  
 A:Experimental source: clone R04D3  
 C:Genetics:  
 A:Gene: CSP:R04D3.3  
 A:Map position: X  
 A:Introns: 108/3; 296/3; 352/3

Query Match 68.2%; Score 30; DB 2; Length 430;  
 Best Local Similarity 57.1%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXGCG 10  
 DB 36 CSPNAGC 42

RESULT 22  
 S01793  
 arylsulfatase (EC 3.1.6.1) precursor - sea urchin (Hemicentrotus pulcherrimus)  
 C:Species: Hemicentrotus pulcherrimus  
 C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1991 #text\_change 05-Oct-2004  
 C:Accession: S01793  
 R:Sasaki, H.; Yamada, K.; Akasaka, K.; Suzuki, K.; Saito, A.; Sato, M.; Sh  
 Eur. J. Biochem. 177, 9-13, 1988  
 A:Title: cDNA cloning, nucleotide sequence and expression of the gene for arylsulfatase  
 A:Reference number: S01793; MUID:89030699; PMID:3181160  
 A:Accession: S01793  
 A:Molecule type: mRNA  
 A:Residues: 1-551 <SAS>  
 A:Cross-references: UNIPROT:P14000; UNIPARC:UPI00001260B5; EMBL:X17015; NID:g9432; PID:g  
 A:Note: part of this sequence was confirmed by protein sequencing  
 A:Note: the authors translated the codon CAC for residue 61 as Asp  
 C:Superfamily: Animal sulfatase  
 C:Keywords: sulfuric ester hydrolase  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:1-551/Product: arylsulfatase #status predicted <MNT>  
 F:100/Modified site: 3-oxoalanine (Cys) #status predicted

Query Match 68.2%; Score 30; DB 2; Length 551;  
 Best Local Similarity 57.1%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXGCG 10  
 DB 518 CNPANGC 524

RESULT 23  
 S07089  
 arylsulfatase (EC 3.1.6.1) - sea urchin (Lytechinus pictus)  
 C:Species: Lytechinus pictus (painted urchin)  
 C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 05-Oct-2004  
 C:Accession: S07089  
 R:Yamada, K.; Akasaka, K.; Shimada, H.  
 Eur. J. Biochem. 186, 405-410, 1989  
 A:Title: structure of sea-urchin arylsulfatase gene.



A:Reference number: S07089; MUID:90092130; PMID:2598936  
A:Accession: S07089  
A:Molecule type: DNA  
A:Residues: 1-551 <YAM>  
A:Cross-references: UNIPROT:P14000; UNIPARC:UPI00001759C4; EMBL:X16679  
C:Genetics:  
A:Gene: Ars  
A:Introns: 103/2; 160/1; 197/3; 272/1; 327/1  
C:Superfamily: Animal sulfatase  
C:Keywords: sulfuric ester hydrolase  
F:100/Modified site: 3-oxoalanine (Cys) #status predicted

Query Match 68.2%; Score 30; DB 2; Length 551;  
Best Local Similarity 57.1%; Pred. No. 2.7e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 518 CNPANGC 524

RESULT 24  
T23864  
hypothetical protein F39D8.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T23864  
R:Hemby, C.  
submitted to the EMBL Data Library, February 1996  
A:Reference number: Z19809  
A:Accession: T23864  
A:Status: preliminary; translated from GB/EMBL/DDbJ  
A:Molecule type: DNA  
A:Residues: 1-648 <MTL>  
A:Cross-references: UNIPROT:Q20191; UNIPARC:UPI000017B99F; EMBL:Z69793; PIDN:CAA93676.1;  
C:Genetics:  
A:Experimental source: clone R03A10  
A:Gene: CESP:F39D8.4  
A:Map position: X  
A:Introns: 22/2; 112/3; 160/1; 250/1; 334/1; 424/3; 463/1; 485/3; 528/2; 593/2

Query Match 68.2%; Score 30; DB 2; Length 648;  
Best Local Similarity 57.1%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 50 CLPTAGC 56

RESULT 25  
E89753  
protein F11C7.4 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: E89753  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: E89753  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1722 <STO>  
A:Cross-references: UNIPROT:Q19350; UNIPARC:UPI0000074975; GB:chr\_X; PIDN:AAC69012.1; PI  
C:Genetics:  
A:Gene: F11C7.4  
A:Map position: X

Query Match 68.2%; Score 30; DB 2; Length 1722;  
Best Local Similarity 57.1%; Pred. No. 6.5e+02;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 401 CEFKSGC 407

RESULT 26  
H71262  
hypothetical protein TP0940 - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C>Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: H71262  
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwinn, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uitterback, T.; McDaniel, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; MUID:98332770; PMID:9665876  
A:Accession: H71262  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-46 <COL>  
A:Cross-references: UNIPROT:O83910; UNIPARC:UPI0000139E45; GB:AE001262; GB:AE000520; NIT  
C:Genetics:  
A:Gene: TP0940

Query Match 65.9%; Score 29; DB 2; Length 46;  
Best Local Similarity 57.1%; Pred. No. 60;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 36 CFPAPGC 42

RESULT 27  
B45937  
early chorion protein Era.1 precursor - silkworm  
C:Species: Bombyx mori (silkworm)  
C>Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 09-Jul-2004  
C:Accession: B45937  
R:Ritner, B.L.; Burke, W.D.; Lecanidou, R.; Rodakis, G.C.; Eichbush, T.H.  
Dev. Biol. 125, 423-431, 1988  
A:Title: Organization and expression of three genes from the silkworm early chorion loci  
A:Reference number: A45937; MUID:88112521; PMID:3338621  
A:Accession: B45937  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-119 <HIB>  
A:Cross-references: UNIPROT:P13511; UNIPARC:UPI0000127783; GB:M19076; NID:G155974; PIDN  
C:Superfamily: chorion class A protein pc292

Query Match 65.9%; Score 29; DB 2; Length 119;  
Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 107 CGPTGCG 113

RESULT 28  
S24292  
chorion protein - silkworm  
C:Species: Bombyx mori (silkworm)  
C>Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C:Accession: S24292  
R:Hibner, B.L.; Burke, W.D.; Eichbush, T.H.  
Genetics 128, 595-606, 1991  
A:Title: Sequence identity in an early chorion multigene family is the result of local  
A:Reference number: S24291; MUID:91340109; PMID:1874417

A:Accession: S24292  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-119 <HIB>  
A:Cross-references: UNIPROT:Q17213; UNIPARC:UPI0000127787; EMBL:X58446; NID:G5852; PIDN:  
C:Genetics:  
A:introns: 17/3  
C:Superfamily: chorion class A protein pc292

Query Match 65.9%; Score 29; DB 2; Length 119;  
Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 107 CGPTGCG 113

RESULT 29  
S24294  
chorion protein - silkworm  
C:Species: Bombyx mori (silkworm)  
C>Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C:Accession: S24294  
R:Hitner, B.L.; Burke, W.D.; Eickbush, T.H.  
Genetics 128, 595-606, 1991  
A>Title: Sequence identity in an early chorion multigene family is the result of localiz  
A:Reference number: S24291; MUID:91340109; PMID:1874417  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-119 <HIB>  
A:Cross-references: UNIPROT:Q17214; UNIPARC:UPI0000127789; EMBL:X58448; NID:G5856; PIDN:  
C:Genetics:  
A:introns: 17/3  
C:Superfamily: chorion class A protein pc292

Query Match 65.9%; Score 29; DB 2; Length 119;  
Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 107 CGPTGCG 113

RESULT 30  
S24291  
chorion protein - silkworm  
C:Species: Bombyx mori (silkworm)  
C>Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C:Accession: S24291  
R:Hitner, B.L.; Burke, W.D.; Eickbush, T.H.  
Genetics 128, 595-606, 1991  
A>Title: Sequence identity in an early chorion multigene family is the result of localiz  
A:Reference number: S24291; MUID:91340109; PMID:1874417  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-119 <HIB>  
A:Cross-references: UNIPROT:Q17212; UNIPARC:UPI0000127785; EMBL:X58445; NID:G5850; PIDN:  
C:Genetics:  
A:introns: 17/3  
C:Superfamily: chorion class A protein pc292

Query Match 65.9%; Score 29; DB 2; Length 119;  
Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 107 CGPTGCG 113

RESULT 31  
S24293  
chorion class CA protein mC11 precursor - silkworm  
C:Species: Bombyx mori (silkworm)  
C>Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 09-Jul-2004  
C:Accession: S24293; A23548  
R:Hitner, B.L.; Burke, W.D.; Eickbush, T.H.  
Genetics 128, 595-606, 1991  
A>Title: Sequence identity in an early chorion multigene family is the result of localiz  
A:Reference number: S24291; MUID:91340109; PMID:1874417  
A:Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-121 <HIB>  
A:Cross-references: UNIPROT:P08829; UNIPARC:UPI0000127788; EMBL:X58447; NID:G5854; PIDN:  
R:Leccandou, R.; Rodakis, G.C.; Eickbush, T.H.; Kafatos, F.C.  
Proc. Natl. Acad. Sci. U.S.A. 83, 6514-6518, 1986  
A>Title: Evolution of the silk moth chorion gene superfamily: gene families CA and CB.  
A:Reference number: A94122; MUID:86313609; PMID:3462711  
A:Accession: A23548  
A:Molecule type: mRNA  
A:Residues: 16-121 <LEC>  
A:Cross-references: UNIPARC:UPI00016888A; GB:M13834; NID:G155969; PIDN:AAA27828.1; PID:  
C:Genetics:  
A:introns: 19/3  
C:Superfamily: chorion class A protein pc292  
C:Keywords: egg

Query Match 65.9%; Score 29; DB 2; Length 121;  
Best Local Similarity 57.1%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 109 CGPTGCG 115

RESULT 32  
E82523  
hypothetical protein XP2732 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: E82523  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A>Note: for a complete list of authors see reference number A59328 below  
A:Accession: E82523  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-147 <SIM>  
A:Cross-references: UNIPROT:Q9P978; UNIPARC:UPI00000C2B2A; GB:AE004079; GB:AE003849; NID:  
A:Experimental source: strain 9a5c  
R:Simpon, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
as-Neto, B.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kempfer, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.B.; Laizy  
Chado, M.A.; Madalira, A.M.B.N.; Madalira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
M.; Teshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XP2732

Query Match 65.9%; Score 29; DB 2; Length 147;

Best Local Similarity 57.1%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 138 CAPKRC 144

## RESULT 33

T33130

Hypothetical protein C23H5.9 - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T33130

R/Author: E. J. Kramer, J.

Submitted to the EMBL Data Library, May 1998

A/Description: The sequence of *C. elegans* cosmid C23H5.

A/Reference number: Z21286

A/Accession: T33130

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-163 &lt;LAM&gt;

A/Cross-references: UNIPROT:O61832; UNIPARC:UPI0000080484; EMBL:AF067609; PIDN:AAC17537.

A/Experimental source: strain Bristol N2; clone C23H5

C/Genetics:

A/Map position: 4

A/Introns: 1/3; 101/3; 126/2

## Query Match

Best Local Similarity 57.1%; Pred. No. 1.6e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 44 CLPRLGC 50

## RESULT 34

P84680

Hypothetical protein At2g28090 [imported] - *Arabidopsis thaliana*C/Species: *Arabidopsis thaliana* (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C/Accession: P84680

R/Author: X. J. Kaul, S. Rounsley, S. D. Shea, T. P. Bent, M. I. Town, C. D. Fujii, C. Y. M.; K. O. H. Molnar, K. S. Cronin, L. A. Shen, M. Y. Vanden, S. E. Umayam, L. Tallon, L. E. D. Nielsen, W. C. White, O. E. Eisen, J. A. Salzberg, S. L. Fraser, C. M. Venter, D. Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: P84680

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-245 &lt;STO&gt;

A/Cross-references: UNIPROT:Q9ZUV1; UNIPARC:UPI000009EFC4; GB:AE002093; NID:G4063749; PI

C/Genetics:

A/Map position: 2

## Query Match

Best Local Similarity 57.1%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 210 CQPGNGC 216

## RESULT 35

P90172

Hypothetical protein thif [imported] - *Sulfolobus solfataricus*C/Species: *Sulfolobus solfataricus*

C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004

C/Accession: P90172

R/Author: O. J. Singh, R. K. Confalonieri, F. Zivanovic, Y. Allard, G. Awey, M. J. Chan-Jong, I. J. Jeffries, A. C. Kozera, C. J. Medina, N. Peng, X. Thi-Ngoc, H. P. Redder, F. arrett, R. A. Ragan, M. A. Senses, C. W. Van der Oost, J.

Submitted to Genbank, April 2001

A/Description: *Sulfolobus solfataricus* complete genome.

A/Reference number: A99139

A/Accession: P90172

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-333 &lt;KUR&gt;

A/Cross-references: UNIPROT:Q980J4; UNIPARC:UPI0000064201; GB:AE006641; NID:G13813443; F

C/Genetics:

A/Map position: 4

A/Introns: 1/3; 101/3; 126/2

## Query Match

Best Local Similarity 57.1%; Pred. No. 2.8e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 321 CYPSTGC 327

## RESULT 36

T35248

probable oxidoreductase - *Streptomyces coelicolor*C/Species: *Streptomyces coelicolor*

C/Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004

C/Accession: T35248

R/Author: K. J. Hartley, D. J. Bentley, S. D. Patkhill, J. J. Barrell, B. G. Rajandream, M. A. submitted to the EMBL Data Library, April 1999

A/Reference number: Z21573

A/Accession: T35248

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-348 &lt;OLI&gt;

A/Cross-references: UNIPROT:Q9X7M9; UNIPARC:UPI00000DAF9E; EMBL:AL049587; PIDN:CAB40673

A/Experimental source: strain A3 (2)

C/Genetics:

A/Map position: 4

A/Introns: 1/3; 101/3; 126/2

## Query Match

Best Local Similarity 57.1%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 328 CYPDLGC 334

## RESULT 37

AB3636

Hypothetical protein PA0063 [imported] - *Pseudomonas aeruginosa* (strain PA01)C/Species: *Pseudomonas aeruginosa*

C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004

C/Accession: AB3636

R/Author: C. K. Pham, X. Q. Erwin, A. L. M. Doguchi, S. D. Warren, P. Hickey, M. J. B. adman, S. J. Yuan, Y. J. Brody, L. L. Coulter, S. N. Folger, K. R. Kae, A. J. Lardig, K. J. Lim, . J. Lory, S. J. Olson, M. V.

Nature 406, 959-964, 2000

A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathc

A/Reference number: A82950; MUID:20437337; PMID:10984043

A/Accession: AB3636

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-375 &lt;STO&gt;

A/Cross-references: UNIPROT:Q91769; UNIPARC:UPI00000C4F10; GB:AE004446; GB:AE004091; NII

A/Experimental source: strain PA01

C/Genetics:

A/Map position: 4

A/Introns: 1/3; 101/3; 126/2

## Query Match

Best Local Similarity 57.1%; Pred. No. 2.9e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Best Local Similarity 57.1%; Pred. No. 3e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 136 CFPDAGC 142

## RESULT 38

S14885

hypothetical protein 8 - yeast (*Pichia angusta*)C/Species: *Pichia angusta*

C/Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004

C/Accession: S14885

R/Krutilina, A.I.; Sergina, S.A.; Tikhomirova, L.P.; Kryukov, V.M.

submitted to the EMBL Data Library, April 1991

A/Description: Nucleotide sequence of *Hansenula polymorpha* DNA region complementing DAK.

A/Reference number: S14854

A/Accession: S14885

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-379 &lt;RRU&gt;

A/Cross-references: UNIPROT:Q04326; UNIPARC:UPI00006AASF; EMBL:X58862; NID:92762; PIND:

Query Match 65.9%; Score 29; DB 2; Length 379;  
Best Local Similarity 57.1%; Pred. No. 3.1e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 372 CRPTTAC 378

## RESULT 39

E96676

hypothetical protein T23K8.15 [imported] - *Arabidopsis thaliana*C/Species: *Arabidopsis thaliana* (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C/Accession: E96676

R/Theologian, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurce, J.S.; Maitl, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: E96676

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-449 &lt;STO&gt;

A/Cross-references: UNIPROT:Q9S9K4; UNIPARC:UPI0000AA2D4; GB:AE05173; NID:94646203; PI

C/Genetics:

A/Map position: 1

A/Map position: 1

Query Match 65.9%; Score 29; DB 2; Length 449;  
Best Local Similarity 57.1%; Pred. No. 3.5e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 148 CQPALGC 154

## RESULT 40

T23375

hypothetical protein K06G5.1 - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T23375

R/Mortimore, B.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z19733

A/Accession: T23375

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-507 &lt;ML&gt;

A/Cross-references: UNIPROT:Q9XUJ9; UNIPARC:UPI000007B186; EMBL:Z81565; PIND:CA04581.1;

A/Experimental source: clone K06G5

C/Genetics:

A/Map position: X

A/Map position: X

A/Introns: 14/1; 71/3; 103/2; 135/1; 199/3; 242/3; 268/2; 306/2; 377/1; 430/3; 461/2

Query Match 65.9%; Score 29; DB 2; Length 507;  
Best Local Similarity 57.1%; Pred. No. 3.8e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 52 CYPDAGC 58

## RESULT 41

D88991

protein apx-1 [imported] - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

C/Accession: D88991

R/anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A/Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology

A/Reference number: A75000; MUID:99069613; PMID:9851916

A/Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele

A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A/Accession: D88991

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-513 &lt;STO&gt;

A/Cross-references: UNIPARC:UPI00017A664; GB:chr\_V; PIND:AAC6353.1; PID:93808343; GSPD

C/Genetics:

A/Map position: 5

A/Map position: 5

Query Match 65.9%; Score 29; DB 2; Length 513;  
Best Local Similarity 57.1%; Pred. No. 3.9e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 207 CUPRAGC 213

## RESULT 42

JC7570

Delta-4 protein - human

C/Species: *Homo sapiens* (man)

C/Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 09-Jul-2004

C/Accession: JC7570

R/Toneya, T.; Tahara, T.; Nagao, K.; Yamada, Y.; Yamamoto, T.; Osawa, M.; Miyatani, S.;

J. Biochem. 129, 27-34, 2001

A/Title: Molecular cloning of Delta-4, a new mouse and human notch ligand.

A/Reference number: JC7569; MUID: 21064937; PMID:11134954

A/Accession: JC7570

A/Molecule type: mRNA

A/Residues: 1-685 &lt;YON&gt;

A/Cross-references: UNIPROT:Q9NR61; UNIPARC:UPI000047825; DDJB:AB043894

C/Comment: This protein, a member of the Notch family of proteins, is a transmembrane re

ates the Notch signaling, the growth or differentiation of vascular endothelial cells.

C/Genetics:

A/Map position: 4

A/Map position: 4

C/Superfamily: delta-4 protein; EGF homology

C:Keywords: transmembrane protein

Query Match 65.9%; Score 29; DB 2; Length 685;  
Best Local Similarity 57.1%; Pred. No. 4.8e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 253 CIPHNCG 259

RESULT 43

JC7569

Delta-4 protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 09-Jul-2004

C:Accession: J07569

C:Yoneda, T.; Tahara, T.; Nagao, K.; Yamada, Y.; Yamamoto, T.; Osawa, M.; Miyatani, S.;

J. Biochem. 129, 27-34, 2001

A:Title: Molecular cloning of Delta-4, a new mouse and human notch ligand.

A:Reference number: JC7569; MUID: 21064937; PMID:11134954

A:Accession: JC7569

A:Molecule type: mRNA

A:Residues: 1-686 <YON>

A:Cross-references: UNIPROT:Q9DBU9; UNIPARC:UPI000000B286; DDBJ:AB043893

C:Comment: This protein, a member of the Notch family of proteins, is a transmembrane re

ates the Notch signaling, the growth or differentiation of vascular endothelial cells.

C:Genetics:

A:Gene: delta-4

C:Superfamily: delta-4 protein; EGF homology

Query Match 65.9%; Score 29; DB 2; Length 686;  
Best Local Similarity 57.1%; Pred. No. 4.8e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 254 CIPHNCG 260

RESULT 44

T28787

hypothetical protein C41D11.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T28787

R:Gatung, S.; Magill, L.

submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid C41D11.

A:Reference number: Z20522

A:Accession: T28787

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-753 <GAT>

A:Cross-references: UNIPARC:UPI000017B826; EMBL:AF003740; PIDD:AA048141.1; GSPDB:GN000019

C:Genetics:

A:Gene: CESP:C41D11.5

A:Map position: 1

A:introns: 53/2; 81/3; 117/1; 250/3; 274/2; 357/3; 443/2; 485/3; 544/3; 585/3; 637/2

Query Match 65.9%; Score 29; DB 2; Length 753;  
Best Local Similarity 57.1%; Pred. No. 5.2e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 338 CEPATNC 344

RESULT 45

A27340

complement C7 precursor [validated] - human

C:Species: Homo sapiens (man)

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: A27340; B53072; C53072

R:Disclipio, R.G.; Chakravarti, D.N.; Muller-Berhard, H.J.; Fey, G.H.

J. Biol. Chem. 263, 549-560, 1988

A:Title: The structure of human complement component C7 and the C5b-7 complex.

A:Reference number: A27340; MUID:88087145; PMID:3335508

A:Accession: A27340

A:Molecule type: mRNA

A:Residues: 1-843 <DIS>

A:Cross-references: UNIPROT:P10643; UNIPARC:UPI00000127C58; GB:J03507; NID:G179715; PIDD:

R:Hobart, M.J.; Fernie, B.; Disclipio, R.G.;

Biochemistry 32, 6198-6205, 1993

A:Title: Structure of the human C6 gene.

A:Reference number: A53072; MUID:93291175; PMID:8512929

A:Accession: B53072

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-15, 'P', 17-21, 'R', 22-143 <HOB1>

A:Cross-references: UNIPARC:UPI00001742CF; GB:X72190; NID:G312799; GB:X72191; NID:G31280

A:Note: sequence extracted from NCBI backbone (NCBIP:134085)

A:Note: this translation is not annotated in Genbank entries HSC7X2, HSC7X3, and HSC7X4,

A:Accession: C53072

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 584-596, 'TY', 599-624, 'N', 626-627, 'L', 629-690, 'L' <HOB2>

A:Cross-references: UNIPARC:UPI00001742D0; GB:X72193; NID:G312797; GB:X72194; NID:G31279

A:Note: sequence extracted from NCBI backbone (NCBIP:134087)

A:Note: this translation is not annotated in Genbank entries HSC7X13, and HSC7X14, relea

R:Hoisteenge, J.; Blommestein, M.; Hess, D.; Furman, A.; Mitroshchenko, O.

J. Biol. Chem. 274, 32786-32794, 1999

A:Title: The four terminal components of the complement system are C-mannosylated on mu

A:Reference number: A59362; MUID:20020247; PMID:10551839

A:Contents: annotation

A:Note: identification and location of C-mannosylation sites by mass-spectroscopy

C:Genetics:

A:Gene: GDB:C7

A:Cross-references: GDB:119046; OMIM:217070

A:Map position: 5p13-5p13

C:Complex: combines with complement C5b-6 complex to form the complement C5b-7 complex

C:Function:

A:Description: as a component of the complement C5b-8 complex, assists in the polymeriza

A:Pathway: complement pathway

C:Superfamily: complement c6; complement factor H repeat homology; EGF homology; LDL rec

C:Keywords: complement alternate pathway; complement pathway; cytolysis; glycoprotein; m

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-843/Product: complement component C7 #status predicted <MAT>

F:26-80/Domain: thrombospondin type 1 repeat homology <THR1>

F:85-119/Domain: LDL receptor ligand-binding repeat homology <LDL>

F:455-486/Domain: EGF homology <EGF>

F:499-546/Domain: thrombospondin type 1 repeat homology <THR2>

F:571-626/Domain: complement factor H repeat homology <FH01>

F:631-688/Domain: complement factor H repeat homology <FH02>

F:631-688/Domain: complement factor H repeat homology (TIP) #status experimental

F:702-754/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:503-506/Modified site: 2'-mannosyl-tryptophan (TIP) (partial) #status experimental

Query Match 65.9%; Score 29; DB 1; Length 843;  
Best Local Similarity 57.1%; Pred. No. 5.7e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
DB 73 CEPTRGC 79

RESULT 46

T16840

hypothetical protein T10B10.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C:Accession: T16840

R:Geisel, C.

submitted to the EMBL Data Library, October 1995  
A:Description: The sequence of C. elegans cosmid T10E10.

A:Reference number: Z18588

A:Accession: T16840

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1101 <GE1>

A:Cross-references: UNIPROT:Q22378; UNIPARC:UPI000017BB8F; EMBL:U39644; NID:G1049339; PI

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:T10E10.4

A:Introns: 93/2; 152/2; 191/3; 209/2; 283/3; 303/1; 399/3; 421/1; 440/1; 465/1; 547/3; 7

Query Match 65.9%; Score 29; DB 2; Length 1101;

Best Local Similarity 57.1%; Pred. No. 7e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 443 CTPGLGC 449

#### RESULT 47

T02766

adhesin wt-1 - Afellowyces dermatitidis

C:Species: Afellowyces dermatitidis

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T02766

R:Hogan, L.H.; Josvai, S.; Klein, B.S.

J. Biol. Chem. 270, 30725-30732, 1995

A:Title: Genomic cloning, characterization, and functional analysis of the major surface

A:Reference number: Z14730; PMID:96107236; PMID:8530512

A:Accession: T02766

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1146 <HOG>

A:Cross-references: UNIPROT:Q00048; UNIPARC:UPI000006C854; EMBL:U37772; NID:G1022895; PI

A:Experimental source: ATCC 26199

Query Match 65.9%; Score 29; DB 2; Length 1146;

Best Local Similarity 57.1%; Pred. No. 7.2e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 1067 CPTTSC 1073

#### RESULT 48

T26859

hypothetical protein Y43F8B.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T26859

R:Rincough, R.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z20278

A:Accession: T26859

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1743 <WIL>

A:Cross-references: UNIPROT:Q9XW5; UNIPARC:UPI0000078FD6; EMBL:AL032623; PIDN:CAA21511.

A:Experimental source: clone Y43F8B

C:Genetics:

A:Gene: CESP:Y43F8B.3

A:Introns: 65/1; 92/2; 128/1; 229/1; 367/1; 422/1; 486/2; 523/1; 571/1; 628/1; 857/2; 94

Query Match 65.9%; Score 29; DB 2; Length 1743;

Best Local Similarity 57.1%; Pred. No. 9.9e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 111

DB 438 CAPGCGC 444

#### RESULT 49

T10053

laminin alpha 5 chain - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004

C:Accession: T10053

R:Miner, J.H.; Lewis, R.M.; Sames, J.R.

submitted to the EMBL Data Library, November 1997

A:Reference number: Z16923

A:Accession: T10053

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-3635 <MIN>

A:Cross-references: UNIPROT:Q61001; UNIPARC:UPI000004C5E8; EMBL:U37501; NID:G2599231; PI

C:Genetics:

A:Gene: Lama5

C:Keywords: basement membrane; cell binding; extracellular matrix

F:1888-1939/Domain: laminin-type EGF-like homology <LEG>

F:1942-1970/Domain: EGF homology <EGF>

Query Match 65.9%; Score 29; DB 2; Length 3635;

Best Local Similarity 57.1%; Pred. No. 1.7e+03;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 1441 CHPLVGC 1447

#### RESULT 50

S25773

testis-specific protein Mst84DB - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004

C:Accession: S25773; B56565

R:Kuhn, R.; Kuhn, C.; Boersch, D.; Glaetzer, K.H.; Schaefer, U.; Schaefer, M.

Mech. Dev. 35, 143-151, 1991

A:Title: A cluster of four genes selectively expressed in the male germ line of Drosophi

A:Reference number: A56565; PMID:92102953; PMID:1684716

A:Accession: S25773

A:Molecule type: DNA

A:Residues: 1-74 <KUH>

A:Cross-references: UNIPROT:Q01643; UNIPARC:UPI000012EAB6; EMBL:X67703; NID:G11072; PIDN

A:Note: sequence extracted from NCBI backbone (NCBIN:74217, NCBI:P.74220)

C:Genetics:

A:Gene: Mst84DB

A:Cross-references: FlyBase:FBgn0004173

A:Map position: 3

C:Superfamily: fruit fly testis-specific protein

C:Keywords: spermatogenesis; tandem repeat

Query Match 63.6%; Score 28; DB 2; Length 74;

Best Local Similarity 57.1%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 68 CREYCGC 74

Search completed: January 4, 2006, 16:09:43  
Job time : 39.4522 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 : Search time 88.9652 Seconds  
(without alignments)  
64.204 Million cell updates/sec

Title: US-09-932-322-1  
Perfect score: 44  
Sequence: 1 XXXCXPXTGCGXXX 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : A\_Geneseq\_21.\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*  
9: geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	81.8	192	8	ADG23148 Bacterial
2	35	79.5	79	4	AAU49413 Propionib
3	35	79.5	79	6	ABM45932 Propionib
4	35	79.5	180	7	ABO74428 Pseudomon
5	35	79.5	371	8	ADT04024 Human pro
6	34	77.3	7	3	AAV62764 PB-cadher
7	34	77.3	7	3	AAV61489 Cadherin-
8	34	77.3	7	3	AAV62007 Cadherin-
9	34	77.3	7	3	AAV62224 Cadherin-
10	34	77.3	7	5	ABU00550 B lymphoc
11	34	77.3	7	5	ABG33862 B lymphoc
12	34	77.3	13	5	ABU00545 B lymphoc
13	34	77.3	13	5	ABG33861 B lymphoc
14	34	77.3	53	4	AAU49332 Propionib
15	34	77.3	53	6	ABM41451 Propionib
16	34	77.3	99	8	ADK87568 Plant full
17	34	77.3	113	6	ADA54961 Human pro
18	34	77.3	138	3	AAAB42027 Human ORF
19	34	77.3	250	7	ABO75329 Pseudomon
20	34	77.3	312	5	ABR93198 Herbicida
21	34	77.3	532	8	ADN17632 Bacterial
22	34	77.3	540	7	ADT05702 Bacterial
23	34	77.3	3226	4	ABG28408 Novel hum
24	33	75.0	158	4	AAU40006 Propionib

25	33	75.0	158	6	ABM36525 Propionib
26	33	75.0	239	8	ADT57626 Arabidops
27	32	72.7	751	3	AAAG3626 Arabidops
28	32	72.7	798	3	AAAG3625 Arabidops
29	32	72.7	835	7	AAAG3624 Arabidops
30	31	70.5	90	7	AAV63264 Arabidops
31	31	70.5	114	5	ABO81251 Pseudomon
32	31	70.5	148	7	ADT71040 Human ORF
33	31	70.5	183	7	ADT58925 Human hea
34	31	70.5	193	4	AAO00015 Human pol
35	31	70.5	193	4	ADT60298 Human con
36	31	70.5	238	6	ABP92529 Orthosomy
37	31	70.5	239	6	ABP93448 Orthosomy
38	31	70.5	239	6	ABP76713 Streptomy
39	31	70.5	285	8	ADK95230 Plant full
40	31	70.5	310	7	ABO74638 Pseudomon
41	31	70.5	354	8	ADU17203 M. tuberc
42	31	70.5	444	7	ABO70819 Pseudomon
43	31	70.5	482	5	ABR93340 Herbicida
44	31	70.5	580	2	AAW32363 Mycobacte
45	31	70.5	580	2	AAW32431 Mycobacte
46	31	70.5	580	2	AAW64303 Mycobacte
47	31	70.5	580	2	AAW81666 M. tuberc
48	31	70.5	580	2	AAV38968 M. tuberc
49	31	70.5	580	2	AAV39105 M. tuberc
50	31	70.5	602	2	AAV04996 Mycobacte
51	31	70.5	602	2	ADU17226 M. tuberc
52	31	70.5	638	8	ADU17226 M. tuberc
53	31	70.5	638	2	AAV04992 Mycobacte
54	31	70.5	804	7	ABO75305 Pseudomon
55	31	70.5	804	7	ABR98398 Streptomy
56	30	68.2	13	9	AAV64260 Cadherin-
57	30	68.2	13	9	ADV77018 Human TAA
58	30	68.2	55	9	ADV77016 Human TRP
59	30	68.2	63	5	ABP33508 Human ORF
60	30	68.2	76	4	ABM16814 Human pro
61	30	68.2	87	4	AAW78641 Human car
62	30	68.2	94	4	AAU22004 Human car
63	30	68.2	94	7	ADK45972 Human car
64	30	68.2	94	8	ADU07390 Human car
65	30	68.2	109	7	ABG04785 Novel hum
66	30	68.2	109	7	ABO74041 Pseudomon
67	30	68.2	121	7	ABO76722 Pseudomon
68	30	68.2	126	7	ABO72980 Pseudomon
69	30	68.2	133	7	ADB99069 LRP5 cons
70	30	68.2	133	7	ADB82537 LRP5 cons
71	30	68.2	133	7	ABO75949 Pseudomon
72	30	68.2	142	4	ADM19798 Protein e
73	30	68.2	175	4	AAV72022 Human Olf
74	30	68.2	175	4	AAV72525 Human OR-
75	30	68.2	187	4	AAU61592 Propionib
76	30	68.2	187	6	ABM58111 Propionib
77	30	68.2	208	8	ADY13040 Plant full
78	30	68.2	208	8	AAV51896 Chitin bi
79	30	68.2	211	2	AAV51897 Chitin bi
80	30	68.2	342	2	AAV42780 Rat neuro
81	30	68.2	345	2	AAV13907 Serpin pr
82	30	68.2	366	4	ABG15722 Novel hum
83	30	68.2	391	4	ABG20618 Novel hum
84	30	68.2	451	9	ABM94980 M. xanthu
85	30	68.2	452	9	ABM94980 M. xanthu
86	30	68.2	459	8	ADT09703 Hyperther
87	30	68.2	459	8	ADT09703 Human pro
88	30	68.2	469	8	ADM90874 Human pha
89	30	68.2	494	8	ADM90873 Human pha
90	30	68.2	547	4	ABG08052 Novel hum
91	30	68.2	575	4	ADK94751 Plant full
92	30	68.2	613	9	ABM90694 M. xanthu
93	30	68.2	638	9	ABG20620 Novel hum
94	30	68.2	672	4	ABG14072 Novel hum
95	30	68.2	685	4	ABG29714 Novel hum
96	30	68.2	698	7	ADM05627 Human pro
97	30	68.2	718	6	ADA33353 Actinobac
			734	8	ADO28637 Human ADA

98	30	68.2	764	5	ABR05596	AbD05596 Human tes	171	29	65.9	258	4	ABG27045	Abg27045 Novel hum
99	30	68.2	767	4	ABG11534	Abj11534 Novel hum	172	29	65.9	263	7	ABO68160	ABO68160 Pseudomon
100	30	68.2	767	7	ADP60558	AdP60558 Human con	173	29	65.9	293	2	AAR28148	AAR28148 Sugar bee
101	30	68.2	789	7	ABR55542	ABR55542 Human pro	174	29	65.9	336	8	ADO63304	ADO63304 Novel hum
102	30	68.2	834	8	ADN47192	AdN47192 Thermococ	175	29	65.9	342	7	ABO63151	ABO63151 Klebsiell
103	30	68.2	995	7	ABM89334	ABM89334 Rice abio	176	29	65.9	359	8	ADY22332	ADY22332 Plant ful
104	30	68.2	2161	2	AAW87724	AAW87724 Rupestris	177	29	65.9	361	4	AAU30935	AAU30935 Novel hum
105	29	65.9	7	2	ABR63928	ABR63928 Lysozyme	178	29	65.9	394	5	ABP69774	ABP69774 Human pol
106	29	65.9	10	4	AAG96135	AAg96135 Human com	179	29	65.9	426	4	ABR65487	ABR65487 Drosophil
107	29	65.9	10	4	AAG96045	AAg96045 Human com	180	29	65.9	436	7	ABO71510	ABO71510 Pseudomon
108	29	65.9	10	4	AAG96075	AAg96075 Human com	181	29	65.9	436	9	ABR15465	ABR15465 Human pol
109	29	65.9	10	4	AAG96083	AAg96083 Human com	182	29	65.9	439	7	ABO73150	ABO73150 Pseudomon
110	29	65.9	10	4	AAG96073	AAg96073 Human com	183	29	65.9	455	5	ABR91454	ABR91454 Herpidea
111	29	65.9	10	4	AAG96081	AAg96081 Human com	184	29	65.9	485	6	ABO01387	ABO01387 Human pro
112	29	65.9	12	3	AA826019	AAb26019 Human IGE	185	29	65.9	485	8	ADN96146	ADN96146 Human NOV
113	29	65.9	12	3	AAU16744	AAu16744 Peptide E	186	29	65.9	489	6	ABO01396	ABO01396 Human pro
114	29	65.9	12	5	ABJ00367	ABj00367 Human IGE	187	29	65.9	489	8	ADN96164	ADN96164 Human NOV
115	29	65.9	18	4	AAW65407	AAW65407 Human bra	188	29	65.9	500	2	AAW94496	AAW94496 Human del
116	29	65.9	19	4	AAW73162	AAW73162 Protease	189	29	65.9	503	3	AAV68936	AAV68936 Amimo aci
117	29	65.9	19	4	AAW73155	AAW73155 Protease	190	29	65.9	504	3	AAV68935	AAV68935 Amimo aci
118	29	65.9	50	4	AAU62562	AAU62562 Proteonib	191	29	65.9	505	3	AAV68918	AAV68918 Amimo aci
119	29	65.9	50	6	ABM59081	ABm59081 Proteonib	192	29	65.9	505	3	AAV68934	AAV68934 Amimo aci
120	29	65.9	50	8	ABO59332	ABo59332 Human gen	193	29	65.9	506	3	AAV68933	AAV68933 Amimo aci
121	29	65.9	53	4	AAU49306	AAu49306 Proteonib	194	29	65.9	506	3	AAV68917	AAV68917 Amimo aci
122	29	65.9	53	4	ABM45825	ABm45825 Proteonib	195	29	65.9	507	3	AAV68932	AAV68932 Amimo aci
123	29	65.9	55	3	AA842800	AAb42800 Human ORF	196	29	65.9	507	3	AAV68916	AAV68916 Amimo aci
124	29	65.9	64	4	ABG16398	ABg16398 Novel hum	197	29	65.9	508	3	AAV68915	AAV68915 Amimo aci
125	29	65.9	71	5	ABP00915	ABp00915 Human ORF	198	29	65.9	508	3	AAV68931	AAV68931 Amimo aci
126	29	65.9	71	7	ABO82044	ABo82044 Pseudomon	199	29	65.9	509	3	AAV68914	AAV68914 Amimo aci
127	29	65.9	85	4	AAW65401	AAW65401 Human bra	200	29	65.9	510	3	AAV68913	AAV68913 Amimo aci
128	29	65.9	85	6	ADA56925	ADA56925 Human sec	201	29	65.9	524	3	AAV68942	AAV68942 Amimo aci
129	29	65.9	85	6	ADA40775	ADA40775 Human sec	202	29	65.9	525	3	AAV68941	AAV68941 Amimo aci
130	29	65.9	85	7	ADC74135	ADC74135 Human sec	203	29	65.9	526	3	AAV68924	AAV68924 Amimo aci
131	29	65.9	86	3	AAV91452	AAV91452 Human sec	204	29	65.9	526	3	AAV68924	AAV68924 Amimo aci
132	29	65.9	86	8	ADL71524	ADl71524 Novel hum	205	29	65.9	527	3	AAV68923	AAV68923 Amimo aci
133	29	65.9	92	4	ABR03747	ABr03747 Novel hum	206	29	65.9	527	3	AAV68937	AAV68937 Amimo aci
134	29	65.9	92	6	ABU13041	ABu13041 Novel hum	207	29	65.9	528	3	AAV68922	AAV68922 Amimo aci
135	29	65.9	92	8	ADJ29067	ADj29067 Human mus	208	29	65.9	528	3	AAV68938	AAV68938 Amimo aci
136	29	65.9	104	3	AAU01070	AAU01070 Human sec	209	29	65.9	528	6	ABR39450	ABR39450 Human bol
137	29	65.9	106	2	AAW48349	AAW48349 Human bre	210	29	65.9	529	3	AAV68921	AAV68921 Amimo aci
138	29	65.9	109	4	ABG26326	ABg26326 Novel hum	211	29	65.9	529	3	AAV68937	AAV68937 Amimo aci
139	29	65.9	112	4	AAU40576	AAu40576 Proteonib	212	29	65.9	530	3	AAV68920	AAV68920 Amimo aci
140	29	65.9	112	6	ABM73095	ABm73095 Proteonib	213	29	65.9	531	3	AAV68919	AAV68919 Amimo aci
141	29	65.9	116	4	AAW73965	AAW73965 Human bon	214	29	65.9	534	5	ABG11485	ABG11485 D. melano
142	29	65.9	116	5	ABG43858	ABg43858 Human pap	215	29	65.9	605	4	ABG14995	ABG14995 Novel hum
143	29	65.9	116	7	ABO80121	ABo80121 Pseudomon	216	29	65.9	621	6	AAE34037	AAE34037 Murine no
144	29	65.9	122	4	AAU51053	AAu51053 Proteonib	217	29	65.9	653	8	ADP88277	ADP88277 Human del
145	29	65.9	122	6	ABM47572	ABm47572 Proteonib	218	29	65.9	659	2	AAW94497	AAW94497 Human del
146	29	65.9	126	2	AAV59997	AAV59997 Human end	219	29	65.9	659	3	AAV68930	AAV68930 Amimo aci
147	29	65.9	129	8	ADP09225	ADP09225 Human pro	220	29	65.9	660	3	AAV68929	AAV68929 Amimo aci
148	29	65.9	134	6	ADP06446	ADp06446 Alloiococ	221	29	65.9	661	3	AAV68928	AAV68928 Amimo aci
149	29	65.9	134	7	ADP58881	ADp58881 Human pol	222	29	65.9	662	3	AAV68927	AAV68927 Amimo aci
150	29	65.9	137	6	ABO01386	ABo01386 Human pro	223	29	65.9	663	3	AAV68926	AAV68926 Amimo aci
151	29	65.9	137	8	ADN96144	ADn96144 Human NOV	224	29	65.9	664	3	AAV68925	AAV68925 Amimo aci
152	29	65.9	142	4	AAU41380	AAu41380 Proteonib	225	29	65.9	677	3	AAV68948	AAV68948 Amimo aci
153	29	65.9	142	6	ABM73899	ABm73899 Proteonib	226	29	65.9	678	3	AAV68947	AAV68947 Amimo aci
154	29	65.9	146	6	ADP06448	ADp06448 Alloiococ	227	29	65.9	679	3	AAV68946	AAV68946 Amimo aci
155	29	65.9	151	8	ADY06118	ADy06118 Plant ful	228	29	65.9	680	3	AAV68945	AAV68945 Amimo aci
156	29	65.9	155	4	ABG26325	ABg26325 Novel hum	229	29	65.9	681	3	AAV68944	AAV68944 Amimo aci
157	29	65.9	159	7	ABO81950	ABo81950 Pseudomon	230	29	65.9	682	3	AAV68943	AAV68943 Amimo aci
158	29	65.9	159	7	ABO78806	ABo78806 Pseudomon	231	29	65.9	685	2	AAW80813	AAW80813 Nucleotid
159	29	65.9	161	4	ABG04542	ABg04542 Novel hum	232	29	65.9	685	2	AAW94507	AAW94507 Human del
160	29	65.9	167	5	ADP41479	ADp41479 Human CD-	233	29	65.9	685	3	AAV76912	AAV76912 Human not
161	29	65.9	174	4	AAU16948	AAu16948 Human nov	234	29	65.9	685	3	AAV68912	AAV68912 Amimo aci
162	29	65.9	188	4	AAU18112	AAu18112 Novel hum	235	29	65.9	685	4	AAU12215	AAU12215 Human pro
163	29	65.9	188	4	AAU17021	AAu17021 Human nov	236	29	65.9	685	5	AAU83698	AAU83698 Human pro
164	29	65.9	188	4	ABR10465	ABR10465 Human cDN	237	29	65.9	685	5	ABR84954	ABR84954 Human ang
165	29	65.9	188	4	AAU19912	AAu19912 Novel hum	238	29	65.9	685	5	ABR95560	ABR95560 Human pro
166	29	65.9	188	5	ABJ05732	ABj05732 Novel hum	239	29	65.9	685	5	ADY31954	ADY31954 Novel hum
167	29	65.9	188	5	ABP67052	ABp67052 Human pol	240	29	65.9	685	6	ABO17659	ABO17659 Novel hum
168	29	65.9	197	7	ABM89228	ABm89228 Rice abio	241	29	65.9	685	6	ABU80845	ABU80845 Human pro
169	29	65.9	220	8	ABO59256	ABo59256 Human gen	242	29	65.9	685	6	ABO33811	ABO33811 Novel hum
170	29	65.9	235	9	ABW95832	ABw95832 M. xanthu	243	29	65.9	685	6	ABO80913	ABO80913 Human pro



244	29	65.9	685	6	ABP97826	Abp97826	Amino aci	317	29	65.9	685	7	ADB22570	Adb22570	Human PRO
245	29	65.9	685	6	ABR39449	Abri39449	Human H4-	318	29	65.9	685	7	ADB23343	Adb23343	Human PRO
246	29	65.9	685	6	ABU66613	Abu66613	Human PRO	319	29	65.9	685	7	ADA92065	Ada92065	Novel hum
247	29	65.9	685	6	ABU55880	Abu55880	Human not	320	29	65.9	685	7	ADB15128	Adb15128	Human PRO
248	29	65.9	685	6	AAE34036	Aae34036	Human not	321	29	65.9	685	7	ADB83704	Adb83704	Novel hum
249	29	65.9	685	6	ABU59694	Abu59694	Novel sec	322	29	65.9	685	7	ADB80810	Adb80810	Novel hum
250	29	65.9	685	6	ABO24884	AbO24884	Human Not	323	29	65.9	685	7	ADB73351	Adb73351	Novel hum
251	29	65.9	685	6	ABP72568	Abp72568	Human Not	324	29	65.9	685	7	ADB38380	Adb38380	Novel hum
252	29	65.9	685	6	ABU82154	Abu82154	Novel hum	325	29	65.9	685	7	ADB78433	Adb78433	Novel hum
253	29	65.9	685	6	ABU66889	Abu66889	Human sec	326	29	65.9	685	7	ADB37828	Adb37828	Novel hum
254	29	65.9	685	6	ADA56067	Ada56067	Novel hum	327	29	65.9	685	7	ADB66300	Adb66300	Novel hum
255	29	65.9	685	6	ADA76038	Ada76038	Human PRO	328	29	65.9	685	7	ADB85081	Adb85081	Human PRO
256	29	65.9	685	6	ABJ72334	Abj72334	Human PRO	329	29	65.9	685	7	ADB89380	Adb89380	Human PRO
257	29	65.9	685	6	ADA18688	Ada18688	Human PRO	330	29	65.9	685	7	ADB90112	Adb90112	Human PRO
258	29	65.9	685	6	ADA61311	Ada61311	Homo sapi	331	29	65.9	685	7	ADB39213	Adb39213	Novel hum
259	29	65.9	685	6	ADB19096	Adb19096	Novel hum	332	29	65.9	685	7	ADB78187	Adb78187	Novel hum
260	29	65.9	685	6	ADB27637	Adb27637	Novel PRO	333	29	65.9	685	7	ADB87253	Adb87253	Human PRO
261	29	65.9	685	6	ADA86116	Ada86116	Novel hum	334	29	65.9	685	7	ADB84835	Adb84835	Human PRO
262	29	65.9	685	6	ADB15680	Adb15680	Human PRO	335	29	65.9	685	7	ADB46836	Adb46836	Novel hum
263	29	65.9	685	6	ADA47466	Ada47466	Human PRO	336	29	65.9	685	7	ADB83950	Adb83950	Novel hum
264	29	65.9	685	6	ADA67261	Ada67261	Human PRO	337	29	65.9	685	7	ADB86443	Adb86443	Human PRO
265	29	65.9	685	6	ADB30268	Adb30268	Human PRO	338	29	65.9	685	7	ADB73105	Adb73105	Novel hum
266	29	65.9	685	6	ADA85564	Ada85564	Novel hum	339	29	65.9	685	7	ADB77048	Adb77048	Novel hum
267	29	65.9	685	6	ADA96776	Ada96776	Human PRO	340	29	65.9	685	7	ADB34205	Adb34205	Human PRO
268	29	65.9	685	6	ADA79080	Ada79080	Human PRO	341	29	65.9	685	7	ADB35309	Adb35309	Human PRO
269	29	65.9	685	6	ADA87219	Ada87219	Novel hum	342	29	65.9	685	7	ADB33653	Adb33653	Human PRO
270	29	65.9	685	6	ADB16421	Adb16421	Human PRO	343	29	65.9	685	7	ADB34757	Adb34757	Human PRO
271	29	65.9	685	6	ADA91513	Ada91513	Novel hum	344	29	65.9	685	7	ADB35681	Adb35681	Human PRO
272	29	65.9	685	6	ADB14576	Adb14576	Human PRO	345	29	65.9	685	7	ADB46256	Adb46256	Novel hum
273	29	65.9	685	6	ADB18537	Adb18537	Novel hum	346	29	65.9	685	7	ADC36943	Adc36943	Human PRO
274	29	65.9	685	6	ADA93752	Ada93752	Human PRO	347	29	65.9	685	7	ADC21933	Adc21933	Human PRO
275	29	65.9	685	6	ADB19648	Adb19648	Novel hum	348	29	65.9	685	7	ADC50129	Adc50129	Novel hum
276	29	65.9	685	6	ADB12960	Adb12960	Human PRO	349	29	65.9	685	7	ADC71676	Adc71676	Novel hum
277	29	65.9	685	6	ABO43192	AbO43192	Novel hum	350	29	65.9	685	7	ADC59655	Adc59655	Novel hum
278	29	65.9	685	6	ADA74214	Ada74214	Human PRO	351	29	65.9	685	7	ADC49964	Adc49964	Novel hum
279	29	65.9	685	6	ADB24447	Adb24447	Human PRO	352	29	65.9	685	7	ADC49163	Adc49163	Novel hum
280	29	65.9	685	6	ADA81971	Ada81971	Human PRO	353	29	65.9	685	7	ADC49680	Adc49680	Novel hum
281	29	65.9	685	6	ADA74934	Ada74934	Human PRO	354	29	65.9	685	7	ADC47541	Adc47541	Novel hum
282	29	65.9	685	6	ADA85012	Ada85012	Novel hum	355	29	65.9	685	7	ADC52662	Adc52662	Novel hum
283	29	65.9	685	6	ADA84460	Ada84460	Novel hum	356	29	65.9	685	7	ADC57016	Adc57016	Novel hum
284	29	65.9	685	6	ADB29716	Adb29716	Human PRO	357	29	65.9	685	7	ADC60207	Adc60207	Novel hum
285	29	65.9	685	6	ADA80244	Ada80244	Human PRO	358	29	65.9	685	7	ADC50682	Adc50682	Novel hum
286	29	65.9	685	6	ADA75486	Ada75486	Human PRO	359	29	65.9	685	7	ADC65209	Adc65209	Human PRO
287	29	65.9	685	6	ADA46771	Ada46771	Human PRO	360	29	65.9	685	7	ADC54307	Adc54307	Novel hum
288	29	65.9	685	6	ADB35007	Adb35007	Human PRO	361	29	65.9	685	7	ADC53268	Adc53268	Novel hum
289	29	65.9	685	6	ADA93183	Ada93183	Human PRO	362	29	65.9	685	7	ADC58791	Adc58791	Novel hum
290	29	65.9	685	6	ADB26533	Adb26533	Human PRO	363	29	65.9	685	7	ADC55669	Adc55669	Novel hum
291	29	65.9	685	6	ADB30820	Adb30820	Human PRO	364	29	65.9	685	7	ADC58239	Adc58239	Novel hum
292	29	65.9	685	6	ABR61827	Abri61827	Human del	365	29	65.9	685	7	ADC47286	Adc47286	Novel hum
293	29	65.9	685	6	ABJ72462	Abj72462	Human PRO	366	29	65.9	685	7	ADC02913	Adc02913	Novel hum
294	29	65.9	685	6	ADA60748	Ada60748	Homo sapi	367	29	65.9	685	7	ADC89905	Adc89905	Novel hum
295	29	65.9	685	6	ADB23895	Adb23895	Human PRO	368	29	65.9	685	7	ADC69324	Adc69324	Human PRO
296	29	65.9	685	6	ADA96224	Ada96224	Human PRO	369	29	65.9	685	7	ADC48213	Adc48213	Human PRO
297	29	65.9	685	6	ADA80796	Ada80796	Human PRO	370	29	65.9	685	7	ADC09742	Adc09742	Human PRO
298	29	65.9	685	6	ADA95672	Ada95672	Human PRO	371	29	65.9	685	7	ADC78161	Adc78161	Novel hum
299	29	65.9	685	6	ADB25981	Adb25981	Human PRO	372	29	65.9	685	7	ADC04317	Adc04317	Novel hum
300	29	65.9	685	6	ADB21466	Adb21466	Novel hum	373	29	65.9	685	7	ADC066396	Adc066396	Novel hum
301	29	65.9	685	6	ABO34357	AbO34357	Human sec	374	29	65.9	685	7	ADC80273	Adc80273	Novel hum
302	29	65.9	685	6	ADA77245	Ada77245	Human PRO	375	29	65.9	685	7	ADC10780	Adc10780	Human PRO
303	29	65.9	685	7	ADB17985	Adb17985	Human PRO	376	29	65.9	685	7	ADC10565	Adc10565	Human PRO
304	29	65.9	685	7	ADA86668	Ada86668	Novel hum	377	29	65.9	685	7	ADC47661	Adc47661	Human PRO
305	29	65.9	685	7	ADA86771	Ada86771	Novel hum	378	29	65.9	685	7	ADC77915	Adc77915	Novel hum
306	29	65.9	685	7	ABR61756	Abri61756	Human del	379	29	65.9	685	7	ADC79721	Adc79721	Novel hum
307	29	65.9	685	7	ADA46159	Ada46159	Novel hum	380	29	65.9	685	7	ADC11525	Adc11525	Human sec
308	29	65.9	685	7	ADB28189	Adb28189	Human PRO	381	29	65.9	685	7	ADC09190	Adc09190	Human PRO
309	29	65.9	685	7	ADB28741	Adb28741	Human PRO	382	29	65.9	685	7	ADC50878	Adc50878	Novel hum
310	29	65.9	685	7	ADA76693	Ada76693	Human PRO	383	29	65.9	685	7	ADC40903	Adc40903	Novel hum
311	29	65.9	685	7	ADA88323	Ada88323	Novel hum	384	29	65.9	685	7	ADC52042	Adc52042	Human PRO
312	29	65.9	685	7	ADA97328	Ada97328	Human PRO	385	29	65.9	685	7	ADC51124	Adc51124	Novel hum
313	29	65.9	685	7	ADB27085	Adb27085	Human PRO	386	29	65.9	685	7	ADC52782	Adc52782	Human PRO
314	29	65.9	685	7	ADB22018	Adb22018	Novel hum	387	29	65.9	685	7	ADC53334	Adc53334	Novel hum
315	29	65.9	685	7	ABJ72164	Abj72164	Human mem	388	29	65.9	685	7	ADC37318	Adc37318	Human sec
316	29	65.9	685	7	ADA66709	Ada66709	Human PRO	389	29	65.9	685	7	ADC51490	Adc51490	Human PRO

390	29	65.9	685	7	ADD02289	Ad002289	Human	PRO	463	29	65.9	685	8	ADE41526	Ad41526	Human	sec
391	29	65.9	685	7	ADD50605	Ad050605	Human	PRO	464	29	65.9	685	8	ADE23065	Ad23065	Human	PRO
392	29	65.9	685	7	ADD01223	Ad011223	Human	PRO	465	29	65.9	685	8	ADE21334	Ad21334	Novel	hum
393	29	65.9	685	7	ADD53905	Ad053905	Novel	hum	466	29	65.9	685	8	ADD77449	Ad77449	Novel	hum
394	29	65.9	685	7	ADD50359	Ad050359	Human	PRO	467	29	65.9	685	8	ADE20596	Ad20596	Novel	hum
395	29	65.9	685	7	ADD51370	Ad051370	Novel	hum	468	29	65.9	685	8	ADD75661	Ad75661	Human	PRO
396	29	65.9	685	7	ADD92222	Ad092222	Human	PRO	469	29	65.9	685	8	ADD74177	Ad74177	Human	PRO
397	29	65.9	685	7	ADD91118	Ad091118	Human	PRO	470	29	65.9	685	8	ADD74423	Ad74423	Human	PRO
398	29	65.9	685	7	ADRE03732	Ad0303732	Human	PRO	471	29	65.9	685	8	ADD76153	Ad76153	Human	PRO
399	29	65.9	685	7	ADRE32029	Ad032029	Novel	hum	472	29	65.9	685	8	ADD85645	Ad85645	Novel	hum
400	29	65.9	685	7	ADRE21961	Ad021961	Human	PRO	473	29	65.9	685	8	ADRE23617	Ad23617	Human	PRO
401	29	65.9	685	7	ADD79185	Ad079185	Human	PRO	474	29	65.9	685	8	ADE24260	Ad24260	Human	PRO
402	29	65.9	685	7	ADBE41721	Ad041721	Human	PRO	475	29	65.9	685	8	ADD87085	Ad87085	Human	PRO
403	29	65.9	685	7	ADBE17538	Ad017538	Human	PRO	476	29	65.9	685	8	ADE05194	Ad05194	Human	PRO
404	29	65.9	685	7	ADDP91670	Ad091670	Human	PRO	477	29	65.9	685	8	ADD75407	Ad75407	Human	PRO
405	29	65.9	685	7	ADBE33133	Ad033133	Novel	hum	478	29	65.9	685	8	ADD76951	Ad76951	Novel	hum
406	29	65.9	685	7	ADBE33685	Ad033685	Novel	hum	479	29	65.9	685	8	ADD866719	Ad866719	Novel	hum
407	29	65.9	685	7	ADD79737	Ad079737	Human	PRO	480	29	65.9	685	8	ADE88951	Ad888951	Human	PRO
408	29	65.9	685	7	ADD92774	Ad092774	Human	PRO	481	29	65.9	685	8	ADBE18090	Ad018090	Human	PRO
409	29	65.9	685	7	ADBE19194	Ad019194	Human	PRO	482	29	65.9	685	8	ADE88399	Ad88399	Human	PRO
410	29	65.9	685	7	ADBE18662	Ad018662	Human	PRO	483	29	65.9	685	8	ADD77695	Ad77695	Novel	hum
411	29	65.9	685	7	ADBE42858	Ad042858	Human	PRO	484	29	65.9	685	8	ADD77941	Ad77941	Novel	hum
412	29	65.9	685	7	ADD95627	Ad095627	Human	PRO	485	29	65.9	685	8	ADD85399	Ad85399	Novel	hum
413	29	65.9	685	7	ADBE22513	Ad022513	Human	PRO	486	29	65.9	685	8	ADD73931	Ad73931	Human	PRO
414	29	65.9	685	7	ADD78631	Ad078631	Human	PRO	487	29	65.9	685	8	ADD74669	Ad74669	Human	PRO
415	29	65.9	685	7	ADBE32581	Ad032581	Novel	hum	488	29	65.9	685	8	ADD77197	Ad77197	Novel	hum
416	29	65.9	685	7	ADBE42273	Ad042273	Human	PRO	489	29	65.9	685	8	ADD85691	Ad85691	Novel	hum
417	29	65.9	685	7	ADD80289	Ad080289	Human	PRO	490	29	65.9	685	8	ADE05540	Ad050540	Human	PRO
418	29	65.9	685	7	ADD89317	Ad089317	Human	PRO	491	29	65.9	685	8	ADD74915	Ad74915	Human	PRO
419	29	65.9	685	7	ADBE40601	Ad040601	Human	PRO	492	29	65.9	685	8	ADE94419	Ad094419	Human	PRO
420	29	65.9	685	7	ADBE04400	Ad004400	Human	PRO	493	29	65.9	685	8	ADE940830	Ad090830	Human	PRO
421	29	65.9	685	7	ADBE92529	Ad092529	Human	PRO	494	29	65.9	685	8	ADE94971	Ad094971	Human	PRO
422	29	65.9	685	7	ABUE63730	Ab063730	Human	Not	495	29	65.9	685	8	ADE93081	Ad093081	Human	PRO
423	29	65.9	685	7	ADGE21238	Ad0221238	Novel	hum	496	29	65.9	685	8	ADFE34662	Adf34662	Human	PRO
424	29	65.9	685	7	ADGE22879	Ad0222879	Novel	hum	497	29	65.9	685	8	ADBE91977	Ad091977	Novel	hum
425	29	65.9	685	7	ADFP97214	Adf97214	Human	PRO	498	29	65.9	685	8	ADE90278	Ad090278	Human	PRO
426	29	65.9	685	7	ABW02792	Abw02792	Human	del	499	29	65.9	685	8	ADE91425	Ad091425	Human	PRO
427	29	65.9	685	7	ABW02786	Abw02786	Human	del	500	29	65.9	685	8	ADG05727	Adg05727	Novel	hum
428	29	65.9	685	7	ABW02793	Abw02793	Human	del	501	29	65.9	685	8	ADG27281	Adg27281	Human	PRO
429	29	65.9	685	7	ABW02789	Abw02789	Human	del	502	29	65.9	685	8	ADG02004	Adg02004	Human	PRO
430	29	65.9	685	7	ABW02794	Abw02794	Human	del	503	29	65.9	685	8	ADG21790	Adg21790	Novel	hum
431	29	65.9	685	7	ABW02791	Abw02791	Human	del	504	29	65.9	685	8	ADG19860	Adg19860	Human	PRO
432	29	65.9	685	7	ADG80278	Adg80278	Human	PRO	505	29	65.9	685	8	ADG19766	Adg19766	Human	PRO
433	29	65.9	685	7	ADG79726	Adg79726	Human	PRO	506	29	65.9	685	8	ADG23983	Adg23983	Novel	hum
434	29	65.9	685	7	ADHE55018	Adh55018	Human	hum	507	29	65.9	685	8	ADFE98337	Adf98337	Human	PRO
435	29	65.9	685	7	ADHE55570	Adh55570	Novel	hum	508	29	65.9	685	8	ADG03168	Adg03168	Human	PRO
436	29	65.9	685	7	ADIE63789	Adi63789	Novel	hum	509	29	65.9	685	8	ADFE8889	Adf8889	Human	PRO
437	29	65.9	685	7	ADIE64738	Adi64738	Novel	hum	510	29	65.9	685	8	ADFE98889	Adf98889	Human	PRO
438	29	65.9	685	7	ADIE63237	Adi63237	Novel	hum	511	29	65.9	685	8	ADG16474	Adg16474	Human	PRO
439	29	65.9	685	7	ADHE81651	Adh81651	Novel	hum	512	29	65.9	685	8	ADG04933	Adg04933	Human	PRO
440	29	65.9	685	7	ADHE81099	Adh81099	Novel	hum	513	29	65.9	685	8	ADG19200	Adg19200	Human	PRO
441	29	65.9	685	7	ADHE82268	Adh82268	Novel	hum	514	29	65.9	685	8	ADG11344	Adg11344	Novel	hum
442	29	65.9	685	7	ADNI15667	Adn15667	Novel	hum	515	29	65.9	685	8	ADG13037	Adg13037	Human	PRO
443	29	65.9	685	7	ADNI16296	Adn16296	Novel	hum	516	29	65.9	685	8	ADG08094	Adg08094	Novel	hum
444	29	65.9	685	7	ADNI5115	Adn15115	Novel	hum	517	29	65.9	685	8	ADG15264	Adg15264	Human	PRO
445	29	65.9	685	7	ADNI4563	Adn14563	Novel	hum	518	29	65.9	685	8	ADG12123	Adg12123	Novel	hum
446	29	65.9	685	8	ADPC48937	Adc48937	Novel	hum	519	29	65.9	685	8	ADFE96662	Adf96662	Human	PRO
447	29	65.9	685	8	ADBE80825	Adc80825	Novel	hum	520	29	65.9	685	8	ADG05847	Adg05847	Human	PRO
448	29	65.9	685	8	ADBE21088	Adc21088	Novel	hum	521	29	65.9	685	8	ADG23431	Adg23431	Novel	hum
449	29	65.9	685	8	ADBE05932	Adc05932	Human	PRO	522	29	65.9	685	8	ADG03720	Adg03720	Novel	hum
450	29	65.9	685	8	ADDF6273	Adf6273	Human	PRO	523	29	65.9	685	8	ADG24621	Adg24621	Novel	hum
451	29	65.9	685	8	ADDF75161	Adf75161	Human	PRO	524	29	65.9	685	8	ADFE94680	Adf94680	Novel	hum
452	29	65.9	685	8	ADDF75907	Adf75907	Novel	hum	525	29	65.9	685	8	ADG06918	Adg06918	Novel	hum
453	29	65.9	685	8	ADDB85139	Adb85139	Novel	hum	526	29	65.9	685	8	ADG07470	Adg07470	Novel	hum
454	29	65.9	685	8	ADDB6965	Adb6965	Novel	hum	527	29	65.9	685	8	ADG06776	Adg06776	Human	PRO
455	29	65.9	685	8	ADDE20842	Adc20842	Novel	hum	528	29	65.9	685	8	ADG54965	Adg54965	Novel	hum
456	29	65.9	685	8	ADBE39139	Adc39139	Novel	hum	529	29	65.9	685	8	ADG60629	Adg60629	Novel	hum
457	29	65.9	685	8	ADDB87637	Adb87637	Human	PRO	530	29	65.9	685	8	ADG61733	Adg61733	Novel	hum
458	29	65.9	685	8	ADDB66041	Adb66041	Human	PRO	531	29	65.9	685	8	ADG81934	Adg81934	Human	PRO
459	29	65.9	685	8	ADBE05666	Adc05666	Human	PRO	532	29	65.9	685	8	ADG57173	Adg57173	Novel	hum
460	29	65.9	685	8	ADDF73671	Adf73671	Human	PRO	533	29	65.9	685	8	ADG56621	Adg56621	Novel	hum
461	29	65.9	685	8	ADFE75489	Adf75489	Human	PRO	534	29	65.9	685	8	ADG55517	Adg55517	Novel	hum
462	29	65.9	685	8	ADDF78511	Adf78511	Novel	hum	535	29	65.9	685	8	ADG58277	Adg58277	Novel	hum

536	29	65.9	685	8	ADG70643	Adg70643	Novel hum	609	29	65.9	843	7	ADF90890	Adf90890	Human hep
537	29	65.9	685	8	ADH39120	Adh39120	Novel hum	610	29	65.9	843	8	ADN03920	Adn03920	Antipsoi
538	29	65.9	685	8	ADG57725	Adg57725	Novel hum	611	29	65.9	843	8	ADG87302	Adg87302	Human com
539	29	65.9	685	8	ADG53309	Adg53309	Novel hum	612	29	65.9	872	7	ADG80879	Adg80879	Novel pro
540	29	65.9	685	8	ADG71195	Adg71195	Novel hum	613	29	65.9	1084	6	ABO01391	Ab001391	Human pro
541	29	65.9	685	8	ADG61382	Adg61382	Human PRO	614	29	65.9	1084	6	ADN96154	Adn96154	Human NOV
542	29	65.9	685	8	ADH30344	Adh30344	Human PRO	615	29	65.9	1086	4	ABE71547	AbE71547	Drosophil
543	29	65.9	685	8	ADH11711	Adh11711	Novel hum	616	29	65.9	1088	6	ABO01397	Ab001397	Human pro
544	29	65.9	685	8	ADG52133	Adg52133	Novel hum	617	29	65.9	1088	6	ADN96166	Adn96166	Human NOV
545	29	65.9	685	8	ADG53861	Adg53861	Novel hum	618	29	65.9	1401	4	ABE61146	AbE61146	Drosophil
546	29	65.9	685	8	ADG80830	Adg80830	Human PRO	619	29	65.9	1423	7	ADO44499	AdO44499	Murine SI
547	29	65.9	685	8	ADG56089	Adg56089	Novel hum	620	29	65.9	1564	7	ADC86801	AdC86801	Human GPC
548	29	65.9	685	8	ADH12335	Adh12335	Novel hum	621	29	65.9	2183	6	ABG76428	AbG76428	Apple etc
549	29	65.9	685	8	ADG61181	Adg61181	Novel hum	622	29	65.9	2743	5	ABE81598	AbE81598	Human lam
550	29	65.9	685	8	ADH28268	Adh28268	Human PRO	623	29	65.9	3597	5	ABO09503	AbO09503	Human lam
551	29	65.9	685	8	ADG54413	Adg54413	Novel hum	624	29	65.9	3597	5	ADO10047	AdO10047	Novel hum
552	29	65.9	685	8	ADG59453	Adg59453	Novel hum	625	29	65.9	3597	5	ABO09501	AbO09501	Human lam
553	29	65.9	685	8	ADH433709	Adh433709	Novel PRO	626	29	65.9	3600	8	ADO10043	AdO10043	Novel hum
554	29	65.9	685	8	ADG34210	Adg34210	Novel hum	627	29	65.9	3635	5	ABE81589	AbE81589	Mouse lam
555	29	65.9	685	8	ADH80877	Adh80877	Human PRO	628	29	65.9	3635	5	AAM50357	Aam50357	Mouse lam
556	29	65.9	685	8	ADH33680	Adh33680	Human PRO	629	29	65.9	3635	8	ADP74034	AdP74034	Murine la
557	29	65.9	685	8	ADH69774	Adh69774	Human PRO	630	29	65.9	3690	8	ADM87254	AdM87254	Human pro
558	29	65.9	685	8	ADG09620	Adg09620	Novel hum	631	29	65.9	3695	5	ABE81588	AbE81588	Human lam
559	29	65.9	685	8	ADH15091	Adh15091	Novel hum	632	29	65.9	3695	5	ADT28030	AdT28030	ECM/CAD pr
560	29	65.9	685	8	ADG08968	Adg08968	Novel hum	633	29	65.9	3695	5	ADN04478	AdN04478	Antipsoi
561	29	65.9	685	8	ADH14423	Adh14423	Novel hum	634	29	65.9	3695	5	ADY25780	AdY25780	MRAC
562	29	65.9	685	8	ADH29935	Adh29935	Novel hum	635	29	65.9	3696	5	AAE17310	AAE17310	Human lam
563	29	65.9	685	8	ADH18018	Adh18018	Novel hum	636	29	65.9	3705	5	AAE17309	AAE17309	Human lam
564	29	65.9	685	8	ADM27332	Adm27332	Novel hum	637	29	65.9	4123	7	ABU62079	AbU62079	Human jel
565	29	65.9	685	8	ADH63299	Adh63299	Novel hum	638	29	65.9	4219	5	ADH48718	AdH48718	NOVI prot
566	29	65.9	685	8	ADH77194	Adh77194	Human PRO	639	29	65.9	4561	4	ABG30203	AbG30203	Novel hum
567	29	65.9	685	8	ADK83054	Adk83054	Human PRO	640	29	65.9	5737	7	ADN95328	AdN95328	Human BEC
568	29	65.9	685	8	ADK66690	Adk66690	Human PRO	641	29	65.9	9222	4	ABG21064	AbG21064	Novel hum
569	29	65.9	685	8	ADK00015	Adk00015	Human Not	642	29	65.9	11	4	AAU04530	AAU04530	VEGF base
570	29	65.9	685	8	ADH65316	Adh65316	Human PRO	643	29	65.9	11	4	ABE45060	AbE45060	Rat album
571	29	65.9	685	8	ADM27452	Adm27452	Human PRO	644	29	65.9	11	8	ADJ73673	AdJ73673	Rat serum
572	29	65.9	685	8	ADM41513	Adm41513	Human del	645	29	65.9	12	8	ADJ25802	AdJ25802	Proline t
573	29	65.9	685	8	ADL73027	Adl73027	Human Not	646	29	65.9	14	9	ADY93363	AdY93363	Peptide 1
574	29	65.9	685	8	ADM42176	Adm42176	Human PRO	647	29	65.9	15	4	ABE45080	AbE45080	Rat album
575	29	65.9	685	8	ADM76013	Adm76013	Human Del	648	29	65.9	15	4	ABE45078	AbE45078	Rat album
576	29	65.9	685	8	ADM28038	Adm28038	Human PRO	649	29	65.9	15	8	ADJ73690	AdJ73690	Rat serum
577	29	65.9	685	8	ADQ14406	Adq14406	Human Del	650	29	65.9	15	8	ADJ73692	AdJ73692	Rat serum
578	29	65.9	685	8	ADR20594	Adr20594	Human Not	651	29	65.9	18	4	ABE45325	AbE45325	Human alb
579	29	65.9	685	8	ADR41796	Adr41796	Human Del	652	29	65.9	18	4	ABE45179	AbE45179	Rat album
580	29	65.9	685	8	ADH95520	Adh95520	Human PRO	653	29	65.9	18	4	ABE45197	AbE45197	Rat album
581	29	65.9	685	8	ADI96072	Adi96072	Novel hum	654	29	65.9	18	4	ABE45172	AbE45172	Rat album
582	29	65.9	685	8	ADR88260	Adr88260	Human Del	655	29	65.9	18	4	ABE45200	AbE45200	Rat album
583	29	65.9	685	8	ADR88276	Adr88276	Mus muscu	656	29	65.9	18	8	ADJ73804	AdJ73804	Rat serum
584	29	65.9	685	8	ADR89382	Adr89382	Human Del	657	29	65.9	18	8	ADJ73768	AdJ73768	Rat serum
585	29	65.9	685	8	ADH75458	Adh75458	Human Del	658	29	65.9	18	8	ADJ73775	AdJ73775	Rat serum
586	29	65.9	685	8	ADG32024	Adg32024	Human hum	659	29	65.9	18	8	ADJ73796	AdJ73796	Rat serum
587	29	65.9	685	8	ADP03008	Adp03008	Human PRO	660	29	65.9	18	8	ADJ73802	AdJ73802	Rat serum
588	29	65.9	685	9	ADX58107	Adx58107	Human Not	661	29	65.9	18	8	ADJ73941	AdJ73941	Human ser
589	29	65.9	685	9	ADH70419	Adh70419	Human Del	662	29	65.9	18	8	ADJ73793	AdJ73793	Rat serum
590	29	65.9	685	9	ADG03059	Adg03059	Human sec	663	29	65.9	20	4	ABE45310	AbE45310	Human alb
591	29	65.9	685	9	ABE13805	AbE13805	Cancer ce	664	29	65.9	20	4	ABE45152	AbE45152	Rat album
592	29	65.9	685	9	ABE77854	AbE77854	Human not	665	29	65.9	20	4	ABE45130	AbE45130	Rat serum
593	29	65.9	686	3	AA776911	Aa776911	Mouse not	666	29	65.9	20	8	ADJ73749	AdJ73749	Rat serum
594	29	65.9	686	3	AA768911	Aa768911	Human not	667	29	65.9	20	8	ADJ73727	AdJ73727	Rat serum
595	29	65.9	686	6	ABU55881	AbU55881	Mouse not	668	29	65.9	25	8	ADJ73927	AdJ73927	Human ser
596	29	65.9	686	7	ABW02798	AbW02798	Mouse del	669	29	65.9	25	8	ADG68099	AdG68099	Dog pancer
597	29	65.9	686	7	ABW02788	AbW02788	Mouse del	670	29	65.9	25	8	ADG68098	AdG68098	pig pancer
598	29	65.9	686	7	ABW02796	AbW02796	Mouse del	671	29	65.9	26	2	AAR95893	Aar95893	Cviduct c
599	29	65.9	686	7	ABW02795	AbW02795	Mouse del	672	29	65.9	33	4	ABE42483	AbE42483	Peptide #
600	29	65.9	686	7	ABW02797	AbW02797	Mouse del	673	29	65.9	33	4	ABM62296	AbM62296	Peptide #
601	29	65.9	686	7	ADH88261	Adh88261	Mus muscu	674	29	65.9	33	4	AAW76186	AAW76186	Human bon
602	29	65.9	711	7	ABE84518	AbE84518	Human dia	675	29	65.9	33	4	AAE63372	AAE63372	Human bra
603	29	65.9	741	7	ABO61594	AbO61594	Klebsiell	676	29	65.9	33	4	ABG57904	AbG57904	Human liv
604	29	65.9	758	7	ABM86855	AbM86855	Rice abio	677	29	65.9	47	2	AA414197	AA414197	Fragment
605	29	65.9	764	3	AA768949	Aa768949	Fusion pr	678	29	65.9	48	7	ADG09186	AdG09186	Novel pro
606	29	65.9	797	8	ABH84517	AbH84517	Human dia	679	29	65.9	48	7	ADG08183	AdG08183	Novel pro
607	29	65.9	802	8	ABE84516	AbE84516	Human dia	680	29	65.9	54	4	AAW19876	AAW19876	Peptide #
608	29	65.9	803	4	ABB62003	AbB62003	Drosophil	681	29	65.9	54	4	ABB39808	AbB39808	Peptide #

682	28	63.6	54	4	AAM33398	Aam33398	Peptide #	755	28	63.6	122	3	AAG03985	Aag03985	Human sec
683	28	63.6	54	4	ABB24422	Abb24422	Protein #	756	28	63.6	122	4	ABB52644	Abb52644	Escherich
684	28	63.6	54	4	AAM73188	Aam73188	Human bon	757	28	63.6	125	8	ADU17067	Adu17067	M. tuberc
685	28	63.6	54	4	AAU46604	Aau46604	Protonib	758	28	63.6	127	4	AAU42623	Aau42623	Protonib
686	28	63.6	54	4	AAM60532	Aam60532	Human bra	759	28	63.6	127	6	AAU39142	Aam39142	Protonib
687	28	63.6	54	4	ABG54904	Abg54904	Human liv	760	28	63.6	138	4	AAU29758	Aau29758	Novel hum
688	28	63.6	54	5	ABG43035	Abg43035	Human pep	761	28	63.6	138	9	AAZ20886	Aea20886	Novel hum
689	28	63.6	54	6	ABM43123	Abm43123	Protonib	762	28	63.6	139	4	AAO10791	Aao10791	Human pol
690	28	63.6	54	6	ABP99771	Abp99771	Human sec	763	28	63.6	142	7	ABO74248	AbO74248	Protonib
691	28	63.6	56	6	ABR01263	Abri01263	Human gen	764	28	63.6	144	7	ABO81545	AbO81545	Pseudomon
692	28	63.6	56	7	ADC20517	Adc20517	Human sec	765	28	63.6	144	8	ADP31453	Adp31453	Human sec
693	28	63.6	57	3	ABAB3084	Abab3084	Human sec	766	28	63.6	149	4	AAE04263	Aae04263	Human gen
694	28	63.6	58	6	ABP77249	Abp77249	N. gonorr	767	28	63.6	154	2	AAV33504	Aay33504	Human unl
695	28	63.6	63	3	AAB51734	Aab51734	Human sec	768	28	63.6	156	7	ABO74152	AbO74152	Pseudomon
696	28	63.6	63	3	AAM58881	Aam58881	Protonib	769	28	63.6	159	4	ABG13130	Abg13130	Novel hum
697	28	63.6	63	6	ABM55400	Abm55400	Protonib	770	28	63.6	164	4	AAE04265	Aae04265	Human gen
698	28	63.6	64	4	ABBI6042	Abbi6042	Human ner	771	28	63.6	164	4	AAE01650	Aae01650	Human gen
699	28	63.6	64	4	AAU57824	Aau57824	Protonib	772	28	63.6	166	8	ADU78647	Adx78647	Plant ful
700	28	63.6	69	6	ABM54343	Abm54343	Protonib	773	28	63.6	166	8	ADU71945	Adx71945	Plant ful
701	28	63.6	71	4	AAU52868	Aau52868	Protonib	774	28	63.6	171	8	ADP30794	Adp30794	Human sec
702	28	63.6	71	6	ABM49387	Abm49387	Protonib	775	28	63.6	171	8	ADP30793	Adp30793	Human sec
703	28	63.6	74	4	ABB66555	Abb66555	Drosophil	776	28	63.6	172	4	AAU39136	Aau39136	Protonib
704	28	63.6	74	4	AAU56073	Aau56073	Protonib	777	28	63.6	172	6	ABM35655	Abm35655	Protonib
705	28	63.6	74	6	ABM52592	Abm52592	Protonib	778	28	63.6	177	2	AAV36634	Aay36634	Fragment
706	28	63.6	74	6	ABP56104	Abp56104	Human IGF	779	28	63.6	177	6	ADA11875	Ada11875	Human nov
707	28	63.6	76	4	AAU42937	Aau42937	Protonib	780	28	63.6	180	7	ABR62395	AbR62395	Polypept
708	28	63.6	76	6	ABM39456	Abm39456	Protonib	781	28	63.6	180	8	ADU58667	Adi58667	Human NOV
709	28	63.6	76	8	ABO56663	AbO56663	Human gen	782	28	63.6	180	8	ADM90897	Adm90897	Human pha
710	28	63.6	79	4	AAU45268	Aau45268	Protonib	783	28	63.6	183	8	ADR66909	Adr66909	Human pro
711	28	63.6	79	6	ABM41787	Abm41787	Protonib	784	28	63.6	183	8	ADR66011	Adr66011	Human pro
712	28	63.6	80	4	AAU49012	Aau49012	Protonib	785	28	63.6	189	4	AAE04268	Aae04268	Human gen
713	28	63.6	80	6	ABM45531	Abm45531	Protonib	786	28	63.6	190	4	AAE01680	Aae01680	Human gen
714	28	63.6	83	4	AAU58896	Aau58896	Protonib	787	28	63.6	190	5	ABG63945	Abg63945	Human alb
715	28	63.6	83	6	ABM55415	Abm55415	Protonib	788	28	63.6	190	8	ADU77210	AdU77210	Albumin f
716	28	63.6	84	4	AAW96592	Aaw96592	Human rep	789	28	63.6	190	9	ADU18898	AdU18898	PRO polyp
717	28	63.6	84	7	ABO84011	AbO84011	Pseudomon	790	28	63.6	191	4	AAE04266	Aae04266	Human gen
718	28	63.6	86	3	AAAB4184	Aaab4184	Human can	791	28	63.6	192	8	ADU12911	AdU12911	Plant ful
719	28	63.6	87	4	AAU67356	Aau67356	Protonib	792	28	63.6	199	2	AAE60316	Aae60316	Sheep LFA
720	28	63.6	87	6	ABM63875	Abm63875	Protonib	793	28	63.6	199	8	ADU17821	AdU17821	Reverse t
721	28	63.6	89	4	AAE04225	Aae04225	Human gen	794	28	63.6	205	2	AAU41495	Aau41495	Fragment
722	28	63.6	89	4	ABBI8046	Abbi8046	Peptide #	795	28	63.6	206	8	ADU20449	AdU20449	Human PRO
723	28	63.6	89	4	AAAM31466	Aaam31466	Peptide #	796	28	63.6	207	2	AAW55459	Aay55459	H. pylori
724	28	63.6	89	4	AAW71185	Aaw71185	Human bon	797	28	63.6	207	4	AAW83636	Aam83636	Human lmm
725	28	63.6	89	4	AAW58678	Aaw58678	Human bra	798	28	63.6	218	2	AAW55202	Aay55202	H. pylori
726	28	63.6	89	4	ABG52887	Abg52887	Human liv	799	28	63.6	218	8	AAV65681	Aay65681	C. elegan
727	28	63.6	89	5	ABG40982	Abg40982	Human pep	800	28	63.6	228	8	ADP30764	Adp30764	Human sec
728	28	63.6	89	5	ABG64494	Abg64494	Human alb	801	28	63.6	228	8	ADU69140	AdU69140	Plant ful
729	28	63.6	89	8	ADU77761	AdU77761	Albumin f	802	28	63.6	233	3	AAE09755	Aae09755	Human sec
730	28	63.6	91	4	AAU64239	Aau64239	Protonib	803	28	63.6	233	3	AAV67291	Aay67291	Human sec
731	28	63.6	91	6	ABM60758	Abm60758	Protonib	804	28	63.6	237	3	AAAB09620	Aaab09620	Insulin-1
732	28	63.6	95	4	ADG27779	Adg27779	Human nov	805	28	63.6	237	4	AAV67293	Aay67293	Insulin-1
733	28	63.6	100	4	AAE04262	Aae04262	Human gen	806	28	63.6	243	7	ADU09307	AdU09307	Plant ful
734	28	63.6	104	3	AAAB27978	Aaab27978	Human sec	807	28	63.6	248	8	ADP30792	Adp30792	Human sec
735	28	63.6	108	2	AAW07693	Aaw07693	Dendroide	808	28	63.6	258	8	ABM84306	Abm84306	Human dia
736	28	63.6	109	7	ADM05568	Adm05568	Human pro	809	28	63.6	258	8	ABM84307	Abm84307	Human dia
737	28	63.6	109	8	AAW07694	Aaw07694	Dendroide	810	28	63.6	258	8	ADU09307	AdU09307	Plant ful
738	28	63.6	109	8	ADP30690	Adp30690	Human sec	811	28	63.6	258	7	ADU01611	AdU01611	Enterohae
739	28	63.6	110	4	ABBI1699	Abbi1699	Drosophil	812	28	63.6	248	8	ADP30662	Adp30662	Human sec
740	28	63.6	110	4	ABBI1699	Abbi1699	Human GAB	813	28	63.6	248	8	ADP31443	Adp31443	Human sec
741	28	63.6	110	6	ABJ19814	Abj19814	Androgen-	814	28	63.6	249	8	ADP31443	Adp31443	Human sec
742	28	63.6	111	4	AAAI4931	Aaai4931	Peptide #	815	28	63.6	250	8	ADM90898	Adm90898	Human pha
743	28	63.6	111	4	ABBI9339	Abbi9339	Protein #	816	28	63.6	254	3	AAV67290	Aay67290	Rat insul
744	28	63.6	111	4	AAW67069	Aaw67069	Human bon	817	28	63.6	255	4	AAV67290	Aay67290	Rat insul
745	28	63.6	111	4	AAW54666	Aaw54666	Human bra	818	28	63.6	257	8	ADO70024	AdO70024	Human col
746	28	63.6	111	4	ABG48733	Abg48733	Human liv	819	28	63.6	258	2	AAE22253	Aae22253	Sequence
747	28	63.6	111	4	AAW02659	Aaw02659	Peptide #	820	28	63.6	258	2	AAE21688	Aae21688	Sequence
748	28	63.6	113	5	ABG36727	Abg36727	Human pep	821	28	63.6	258	2	AAW37466	Aay37466	Inhibitor
749	28	63.6	113	3	ABBS6144	Abbs6144	Human sec	822	28	63.6	258	3	AAV67292	Aay67292	Insulin-1
750	28	63.6	116	4	ABG19371	Abg19371	Novel hum	823	28	63.6	258	3	AAV67292	Aay67292	Insulin-1
751	28	63.6	116	8	ADP30955	Adp30955	Human sec	824	28	63.6	258	4	AAW50976	Aaw50976	Human PRO
752	28	63.6	117	6	AAU61137	Aau61137	Protonib	825	28	63.6	258	4	AAW50991	Aaw50991	Human PRO
753	28	63.6	117	6	ABM57656	Abm57656	Protonib	826	28	63.6	258	4	AAW50913	Aaw50913	Human PRO
754	28	63.6	117	6	AAO16319	Aao16319	Human ins	827	28	63.6	258	5	AAU86154	Aau86154	Human PRO

828	28	63.6	258	6	ABU71442	Abu71442 Human neo	901.	28	63.6	410	4	AAE00849	Aae00849 Human nov
829	28	63.6	258	7	ABR62394	Abri62394 Polypepti	902	28	63.6	423	8	ADP30819	Adp30819 Human sec
830	28	63.6	258	7	ADJ37335	Adj37335 Human tum	903	28	63.6	429	8	ADP30570	Adp30570 Human sec
831	28	63.6	258	8	ADH13198	Adh13198 Human mal	904	28	63.6	437	2	AAW67722	Aaw67722 Human tum
832	28	63.6	258	8	ADG68259	Adg68259 Human mal	905	28	63.6	437	3	AAI87750	Aai87750 Human prv
833	28	63.6	258	8	ADJ58665	Adj58665 Human NOV	906	28	63.6	437	3	AAI66737	AAi66737 Membrane-
834	28	63.6	258	8	ADP30597	Adp30597 Human sec	907	28	63.6	437	3	AAI96733	AAi96733 PRO1863,
835	28	63.6	258	8	ADY14554	Ady14554 PRO polyP	908	28	63.6	437	3	AAI87261	AAi87261 Human sig
836	28	63.6	258	9	ADY19576	Ady19576 PRO polyP	909	28	63.6	437	3	AAU12404	AAu12404 Human PRO
837	28	63.6	258	9	ADZ09729	Adz09729 Human bre	910	28	63.6	437	4	AAI70851	AAi70851 Human prv
838	28	63.6	258	9	AEA15085	Aea15085 Human pol	911	28	63.6	437	4	AAE65260	AAe65260 Human PRO
839	28	63.6	260	8	ADK76813	Adk76813 Plant ful	912	28	63.6	437	5	AAE80767	AAe80767 Human gra
840	28	63.6	261	8	ADK67963	Adk67963 Plant ful	913	28	63.6	437	5	AAU83672	AAu83672 Human PRO
841	28	63.6	263	2	AAE60474	Aae60474 E. plinat	914	28	63.6	437	5	ABG92707	ABg92707 Human sec
842	28	63.6	265	2	AEA19937	Aea19937 Novel hum	915	28	63.6	437	5	ABG91359	ABg91359 Novel hum
843	28	63.6	268	4	ABE66405	Abbe66405 Drosophi	916	28	63.6	437	5	ABG31399	ABg31399 Human PRO
844	28	63.6	270	4	ABG28002	Abg28002 Novel hum	917	28	63.6	437	5	ADY31902	ADy31902 Novel hum
845	28	63.6	271	7	ADF42690	Adf42690 S. xenoph	918	28	63.6	437	6	ABU72375	ABu72375 Novel hum
846	28	63.6	273	2	AAI50557	Aai50557 HEV1 IAT	919	28	63.6	437	6	ABU58075	ABu58075 Human PRO
847	28	63.6	276	8	ADK73340	Adk73340 Plant ful	920	28	63.6	437	6	ABU59153	ABu59153 Novel hum
848	28	63.6	277	8	ADP30503	Adp30503 Human sec	921	28	63.6	437	6	ABU80865	ABu80865 Human sec
849	28	63.6	277	8	ADK67533	Adk67533 Plant ful	922	28	63.6	437	6	ABU82665	ABu82665 Human sec
850	28	63.6	279	8	ADP31509	Adp31509 Human sec	923	28	63.6	437	6	ABO17848	ABo17848 Novel hum
851	28	63.6	279	8	ADP31523	Adp31523 Human sec	924	28	63.6	437	6	ABU60584	ABu60584 Human sec
852	28	63.6	279	8	ADP31503	Adp31503 Human sec	925	28	63.6	437	6	ABU80819	ABu80819 Human PRO
853	28	63.6	280	4	ABG04786	Abg04786 Novel hum	926	28	63.6	437	6	ABO33785	ABo33785 Novel hum
854	28	63.6	280	4	ADK77420	Adk77420 Plant ful	927	28	63.6	437	6	ABU13966	ABu13966 Human PRO
855	28	63.6	282	8	ADK77654	Adk77654 Plant ful	928	28	63.6	437	6	ABU81102	ABu81102 Human PRO
856	28	63.6	295	7	ABO69521	AbO69521 Pseudomon	929	28	63.6	437	6	ABU72551	ABu72551 Novel hum
857	28	63.6	296	6	ABP97797	Abp97797 Amilo act	930	28	63.6	437	6	ABU66802	ABu66802 Human PRO
858	28	63.6	297	3	AAI33201	Aai33201 Zea mays	931	28	63.6	437	6	ABG73312	ABg73312 Human PRO
859	28	63.6	300	8	ADP30805	Adp30805 Human sec	932	28	63.6	437	6	ABU59883	ABu59883 Novel sec
860	28	63.6	302	7	ABR61804	Abri61804 Human JAG	933	28	63.6	437	6	ABU59300	ABu59300 Human sec
861	28	63.6	302	7	ABR61734	Abri61734 Human JAG	934	28	63.6	437	6	ABO25997	ABo25997 Human PRO
862	28	63.6	303	4	ABG19635	Abg19635 Novel hum	935	28	63.6	437	6	ABO25073	ABo25073 Human sec
863	28	63.6	304	8	ADP31659	Adp31659 Human sec	936	28	63.6	437	6	ABU82128	ABu82128 Novel hum
864	28	63.6	306	8	ADP31205	Adp31205 Human sec	937	28	63.6	437	6	ABU59006	ABu59006 Human sec
865	28	63.6	309	6	ABU27667	Abu27667 Protein e	938	28	63.6	437	6	ABU60811	ABu60811 Human sec
866	28	63.6	312	8	ADP31454	Adp31454 Human sec	939	28	63.6	437	6	ABU92384	ABu92384 Novel hum
867	28	63.6	315	8	ADP31685	Adp31685 Human sec	940	28	63.6	437	6	ABU59449	ABu59449 Novel hum
868	28	63.6	326	6	ABG99903	Abg99903 Human nov	941	28	63.6	437	6	ABU67078	ABu67078 Human sec
869	28	63.6	327	8	ADP30478	Adp30478 Human sec	942	28	63.6	437	6	ABU81234	ABu81234 Human PRO
870	28	63.6	328	6	ADB09578	AdB09578 Alloiococ	943	28	63.6	437	6	ABU92215	ABu92215 Novel hum
871	28	63.6	331	6	ADB09580	AbD09580 Alloiococ	944	28	63.6	437	6	ABU10921	ABu10921 Human PRO
872	28	63.6	331	7	ABO70490	AbO70490 Pseudomon	945	28	63.6	437	6	ABU81673	ABu81673 Novel hum
873	28	63.6	337	3	AAO70589	Aao70589 Protein e	946	28	63.6	437	6	ABU86512	ABu86512 Human sec
874	28	63.6	338	7	ABO63655	AbO63655 Klebsiell	947	28	63.6	437	6	ABO34126	ABo34126 Human PRO
875	28	63.6	342	1	AAI91375	Aai91375 Vaccinia	948	28	63.6	437	6	ADA45985	Ada45985 Novel hum
876	28	63.6	348	8	ADP31374	Adp31374 Human sec	949	28	63.6	437	6	ADA76416	Ada76416 Human PRO
877	28	63.6	349	5	ABB08605	Abb08605 Hlrudin P	950	28	63.6	437	6	ABU72308	ABu72308 Human PRO
878	28	63.6	350	8	ADP31672	Adp31672 Human sec	951	28	63.6	437	6	ADA19066	Ada19066 Human PRO
879	28	63.6	354	6	ADA57025	Ada57025 Human sec	952	28	63.6	437	6	ADA61689	Ada61689 Homo sapi
880	28	63.6	354	6	ADK40876	Adk40876 Human sec	953	28	63.6	437	6	ADBI9474	ADbi9474 Novel hum
881	28	63.6	354	6	ABR47807	Abri47807 Human sec	954	28	63.6	437	6	ADB28015	ADb28015 Human PRO
882	28	63.6	354	7	ADK74192	AdK74192 Human sec	955	28	63.6	437	6	ADA86494	Ada86494 Novel hum
883	28	63.6	355	2	AAV36267	Aav36267 Human sec	956	28	63.6	437	6	ADBI6058	ADbi6058 Human PRO
884	28	63.6	355	6	ADA11615	Ada11615 Human nov	957	28	63.6	437	6	ADA37866	Ada37866 Human sec
885	28	63.6	357	8	ADP31223	Adp31223 Human sec	958	28	63.6	437	6	ADA47844	Ada47844 Human PRO
886	28	63.6	357	8	ADP30505	Adp30505 Human sec	959	28	63.6	437	6	ADA21552	Ada21552 Human sec
887	28	63.6	360	5	ABP69663	Abp69663 Human sec	960	28	63.6	437	6	ADA10339	Ada10339 Human PRO
888	28	63.6	360	8	ADN04946	Adn04946 Antipsoi	961	28	63.6	437	6	ADA67639	Ada67639 Human PRO
889	28	63.6	360	8	ADK66522	AdK66522 Plant ful	962	28	63.6	437	6	ADB30646	ADb30646 Human PRO
890	28	63.6	363	4	ABG07807	Abg07807 Novel hum	963	28	63.6	437	6	ADA85942	Ada85942 Novel hum
891	28	63.6	366	8	ADP31139	Adp31139 Human sec	964	28	63.6	437	6	ADA11783	Ada11783 Human PRO
892	28	63.6	371	8	ADP31612	Adp31612 Human sec	965	28	63.6	437	6	ADA97154	Ada97154 Human PRO
893	28	63.6	378	8	ADP31375	Adp31375 Human sec	966	28	63.6	437	6	ADA79458	Ada79458 Human PRO
894	28	63.6	384	8	ADP31575	Adp31575 Human sec	967	28	63.6	437	6	ADA87597	Ada87597 Novel hum
895	28	63.6	385	8	ADT57543	Adt57543 Plant pol	968	28	63.6	437	6	ADBI6799	ADbi6799 Human PRO
896	28	63.6	387	7	ADC87455	Adc87455 Human GPC	969	28	63.6	437	6	ADA27991	Ada27991 Human sec
897	28	63.6	387	7	ADP31308	Adp31308 Human sec	970	28	63.6	437	6	ADA91891	Ada91891 Novel hum
898	28	63.6	397	4	AAE00850	Aae00850 Human nov	971	28	63.6	437	6	ADBI4954	ADbi4954 Human PRO
899	28	63.6	405	8	ADP31534	Adp31534 Human sec	972	28	63.6	437	6	ADBI8915	ADbi8915 Novel hum
900	28	63.6	407	8	ADY08554	Ady08554 Plant ful	973	28	63.6	437	6	ADA94130	Ada94130 Human PRO

774 28 63.6 437 6 ADB20026 Novel hum  
775 28 63.6 437 6 ADB13338 Human PRO  
776 28 63.6 437 6 ABO43381 Novel hum  
777 28 63.6 437 6 ADA94571 Human sec  
778 28 63.6 437 6 ADA74592 Human PRO  
779 28 63.6 437 6 ADB24825 Human PRO  
780 28 63.6 437 6 ADA82349 Human PRO  
781 28 63.6 437 6 ADA75312 Human PRO  
782 28 63.6 437 6 ADA85330 Novel hum  
783 28 63.6 437 6 ADA84838 Novel hum  
784 28 63.6 437 6 ADB30094 Human PRO  
785 28 63.6 437 6 ADA80622 Human PRO  
786 28 63.6 437 6 ADA75864 Human PRO  
787 28 63.6 437 6 ADA38796 Human sec  
788 28 63.6 437 6 ADA47089 Human PRO  
789 28 63.6 437 6 ADB25385 Human PRO  
790 28 63.6 437 6 ADA93561 Human PRO  
791 28 63.6 437 6 ADB26911 Human PRO  
792 28 63.6 437 6 ADB31198 Human PRO  
793 28 63.6 437 6 ABB72436 Human PRO  
794 28 63.6 437 6 ABB62955 Human PRO  
795 28 63.6 437 6 ADA92917 Human sec  
796 28 63.6 437 6 ADA61126 Homo sap1  
797 28 63.6 437 6 ADB24273 Human PRO  
798 28 63.6 437 6 ADA96602 Human PRO  
799 28 63.6 437 6 ADA81174 Human PRO  
1000 28 63.6 437 6 ADA96050 Human PRO

## ALIGNMENTS

RESULT 1  
ID ADS23148 standard; protein; 192 AA.  
XX  
AC ADS23148;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polypeptide #12181.  
XX  
KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
OS Bacteria.  
XX  
PN US2003233675-A1.  
XX  
PD 18-DEC-2003.  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
PR 21-FEB-2002; 2002US-0360039P.  
XX  
PA (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX  
DR WPI; 2004-061375/06.  
XX  
PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.

XX  
PS Claim 1; SEQ ID NO 12181; 122pp; English.  
XX  
CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 192 AA;

Query Match 81.8%; Score 36; DB 8; Length 192;  
Best Local Similarity 71.4%; Pred. No. 91;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGCC 10  
Db 176 CSPATGCC 182

RESULT 2  
ID AAU49413 standard; protein; 79 AA.  
XX  
AC AAU49413;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Propionibacterium acnes immunogenic protein #10309.  
XX  
KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
OS Propionibacterium acnes.  
XX  
PN MO200181581-A2.  
XX  
PD 01-NOV-2001.  
XX  
PF 20-APR-2001; 2001WO-US012865.  
XX  
PR 21-APR-2000; 2000US-0199047P.  
PR 02-JUN-2000; 2000US-0208841P.  
PR 07-JUL-2000; 2000US-0216747P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
XX  
DR L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
DR WPI; 2001-616774/71.  
XX  
PT N-PSDB; AAS59545.

XX Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris.

XX Example 1; SEQ ID NO 10608; 10699p; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 79 AA;

Query Match 79.5%; Score 35; DB 4; Length 79;  
Best Local Similarity 71.4%; Pred. No. 67;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
| | | | |  
Db 3 CSPGTGC 9

RESULT 3

ID ABM45932 standard; protein; 79 AA.

XX ABM45932;

DT 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #10608.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;

KW immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

PN WO200303515-A1.

XX 24-APR-2003.

PD 11-OCT-2002; 2002WO-US032727.

PR 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL,  
PI Zhang Y, Wang S, Jen S, Lodes MC, Benson DR, Jones R, Carter D,  
PI Barch B, Vallee-Douglas J;

XX WPI: 2003-381789/36.

DR N-PSDB; ACR64474.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the  
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a P. acnes protein.  
XX Example 1; SEQ ID NO 10608; 1481p; English.

XX The invention relates to an isolated polynucleotide (ACR64435-ACR64733)  
CC encoding a Propionibacterium acnes protein. The invention also relates to  
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
CC immunogenic fragments of P. acnes polypeptides. The invention  
CC additionally encompasses expression vectors and host cells comprising a  
CC polynucleotide of the invention; antibodies against polypeptides of the  
CC invention; fusion proteins comprising a polypeptide of the invention; a  
CC method for stimulating an immune response specific for a P. acnes  
CC polypeptide and an isolated T cell population comprising T cells prepared  
CC via this method; a vaccine composition (comprising P. acnes polypeptides,  
CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
CC antigen-presenting cells that express the polypeptide); a method and kit  
CC for detecting or determining the presence or absence of P. acnes in a  
CC patient; and a method for inhibiting the development of P. acnes in a  
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
CC proteins, T cell populations or antigen-presenting cells that express the  
CC polypeptides are useful for diagnosing, preventing or treating acne  
CC vulgaris, or for stimulating an immune response specific for a P. acnes  
CC protein. The polynucleotides can also be used as probes or primers for  
CC nucleic acid hybridisation. The vaccine composition is useful for the  
CC stimulation of an immune response against P. acnes, or for treating acne,  
CC and the kit is useful for performing a diagnostic assay. The present  
CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
CC reading frame) contained within the P. acnes polynucleotides of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 79 AA;

Query Match 79.5%; Score 35; DB 6; Length 79;  
Best Local Similarity 71.4%; Pred. No. 67;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
| | | | |  
Db 3 CSPGTGC 9

RESULT 4

ID ABO74428 standard; protein; 180 AA.

XX ABO74428;

DT 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #6603.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US6551795-B1.

XX 22-APR-2003.

PD 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfeld MJ, Nolling J, Deloughery C, Bush D;

DR WPI: 2003-615309/58.

XX N-PSDB; ABD07999.



PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
XX  
PS Disclosure; SEQ ID NO 23174; 455bp; English.  
XX  
XX  
CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-*P. aeruginosa* drugs, as templates for recombinant  
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
CC of *Pseudomonas* species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 180 AA;  
XX  
XX  
Query Match 79.5%; Score 35; DB 7; Length 180;  
Best Local Similarity 71.4%; Pred. No. 1.3e+02; Indels 0; Gaps 0;  
Matches 5; Conservative 0; Mismatches 2;  
XX  
QY 4 CXPXTGC 10  
| | | | |  
Db 124 CRPATGC 130  
XX  
XX  
RESULT 5  
ADT04024  
ID ADT04024 standard; protein; 371 AA.  
XX  
XX  
AC ADT04024;  
XX  
DT 30-DEC-2004 (first entry)  
XX  
DE Human protein modification and maintenance molecule protein SeqID4.  
XX  
KW protein modification and maintenance molecule; PMMW; cytostatic;  
KW immunomodulator; expression; immune disorder; cancer; human.  
XX  
OS Homo sapiens.  
XX  
PN WO2004084806-A2.  
XX  
PD 07-OCT-2004.  
XX  
PF 16-MAR-2004; 2004WO-US08006.  
XX  
XX  
PR 21-MAR-2003; 2003US-0456864P.  
PR 03-APR-2003; 2003US-0460512P.  
PR 19-MAY-2003; 2003US-0472827P.  
PR 30-MAY-2003; 2003US-0475072P.  
PR 30-MAY-2003; 2003US-0475190P.  
PR 24-JUN-2003; 2003US-0482575P.  
PR 13-AUG-2003; 2003US-0495151P.  
XX  
XX  
PA (INCY-) INCYTE CORP.  
XX  
XX  
PI Baughn MR, Marquis JP, Kable AE, Chawla NK, Emerling BM, Lee SY,  
PI Hafatia AJA, Ramkumar J, Richardson TW, Wang JT, Nakamura IL,  
PI Yang J, Jin P, Becha SD, Wilson AD;  
XX  
DR WPI; 2004-710256/69.  
DR N-PSDB; ADT04052.  
XX  
PT New human protein modification and maintenance molecules (PMMW)

PT polypeptide, useful in preparing a composition for treating a disease  
PT associated with decreased expression or overexpression of PMMW e.g.,  
PT cancer.  
XX  
XX  
PS Claim 1; SEQ ID NO 4; 222bp; English.  
XX  
XX  
CC This invention relates to novel human protein modification and  
CC maintenance molecule (PMMW) proteins and the DNA sequences which encode  
CC them. The invention may be useful for the production of compounds with a  
CC cytostatic or immunomodulator activity. The proteins are useful in  
CC preparing a composition for diagnosing or treating a disease or condition  
CC associated with decreased expression or overexpression of functional  
CC PMMW, for example immune disorders or cancer. The present sequence is  
CC that of a human protein modification and maintenance molecule (PMMW)  
CC protein of the invention.  
XX  
SQ Sequence 371 AA;  
XX  
XX  
Query Match 79.5%; Score 35; DB 8; Length 371;  
Best Local Similarity 71.4%; Pred. No. 2.4e+02; Indels 0; Gaps 0;  
Matches 5; Conservative 0; Mismatches 2;  
XX  
QY 4 CXPXTGC 10  
| | | | |  
Db 364 CSPHTGC 370  
XX  
XX  
RESULT 6  
AAV62764  
ID AAV62764 standard; peptide; 7 AA.  
XX  
XX  
AC AAV62764;  
XX  
DT 02-MAR-2000 (first entry)  
XX  
DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:4047.  
XX  
KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
KW neurological disease; cyclic.  
XX  
XX  
OS Synthetic.  
XX  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT Disulfide-bond 1..7  
XX  
PN WO9957149-A2.  
XX  
XX  
PD 11-NOV-1999.  
XX  
XX  
PF 05-MAY-1999; 99WO-CA000363.  
XX  
XX  
PR 05-MAY-1998; 98US-00073040.  
PR 06-NOV-1998; 98US-00187859.  
PR 20-JAN-1999; 99US-00234395.  
PR 08-MAR-1999; 99US-00264516.  
XX  
XX  
PA (ADHE-) ADHEREX TECHNOLOGIES INC.  
XX  
XX  
PI Blaschuk OW, Gour BJ, Byers S;  
XX  
DR WPI; 2000-038791/03.  
XX  
XX  
PT New cadherin modulating agents, used for modulating nonclassical cadherin  
PT mediated functions for treating e.g. cancers, obesity, rheumatoid  
PT arthritis, multiple sclerosis, diabetes or a neurological disease.  
XX



PS Claim 72; Page 193; 252pp; English.

XX The present invention describes cadherin modulating agents (MA)

CC comprising peptides which comprise a nonclassical cadherin cell adhesion

CC recognition (CAR) sequence. The MA can be used for modulating

CC nonclassical cadherin-mediated functions. They can be used for e.g.

CC inhibiting adhesion of nonclassical-cadherin expressing cells in a

CC mammal, enhancing delivery of a drug through the skin of a mammal,

CC enhancing delivery of a drug to a tumor in a mammal, treating cancer in

CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting

CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-

CC expressing cell, preventing or treating obesity in a mammal, stimulating

CC blood vessel regression in a mammal, enhancing drug delivery to the

CC central nervous system, treating a demyelinating neurological disease,

CC increasing vasopermeability in a mammal, enhancing adhesion of

CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in

CC a mammal, or preventing pregnancy in a mammal. They can also be used for

CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing

CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a

CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age

CC -related macular degeneration, multiple sclerosis and diabetes. The

CC products can also be used for detection and diagnosis and in bioreactors.

CC AAY60592 to AAY64572 represent specifically claimed peptides, and

CC AAY64573 to AAY64643 and AA233183 to AA233186 represent sequences used in

CC the exemplification of the present invention

XX

SO Sequence 7 AA:

Query Match 77.3%; Score 34; DB 3; Length 7;

Best Local Similarity 71.4%; Pred. No. 2e+06;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 1 CDPXTGC 7

RESULT 7

AAV61489

ID AAY61489 standard; peptide; 7 AA.

XX

AC AAY61489;

XX

DT 02-MAR-2000 (first entry)

XX

DE Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1375.

XX

KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;

KW inhibition; cadherin extracellular domain; cell adhesion recognition;

KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;

KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;

KW cadherin related neuronal receptor; LI-cadherin; protocadherin;

KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;

KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;

KW neurological disease; cyclic.

XX

OS Synthetic.

OS Homo sapiens.

XX

FT Key Location/Qualifiers

FT Disulfide-bond 1..7

XX

PN WO9957149-A2.

XX

PD 11-NOV-1999.

XX

PF 05-MAY-1999; 99WO-CA000363.

XX

PR 05-MAY-1998; 98US-00073040.

PR 06-NOV-1998; 98US-00187859.

PR 20-JAN-1999; 99US-00234395.

PR 08-MAR-1999; 99US-00264516.

XX

PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX

PI Blaschuk OM, Gour BJ, Byers S;

XX

DR WPI, 2000-038791/03.

XX

PT New cadherin modulating agents, used for modulating nonclassical cadherin

XX mediated functions for treating e.g. cancers, obesity, rheumatoid

XX arthritis, multiple sclerosis, diabetes or a neurological disease.

XX

PS Claim 36; Page 172; 252pp; English.

XX

XX The present invention describes cadherin modulating agents (MA)

CC comprising peptides which comprise a nonclassical cadherin cell adhesion

CC recognition (CAR) sequence. The MA can be used for modulating

CC nonclassical cadherin-mediated functions. They can be used for e.g.

CC inhibiting adhesion of nonclassical-cadherin expressing cells in a

CC mammal, enhancing delivery of a drug through the skin of a mammal,

CC enhancing delivery of a drug to a tumor in a mammal, treating cancer in

CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting

CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-

CC expressing cell, preventing or treating obesity in a mammal, stimulating

CC blood vessel regression in a mammal, enhancing drug delivery to the

CC central nervous system, treating a demyelinating neurological disease,

CC increasing vasopermeability in a mammal, enhancing adhesion of

CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in

CC a mammal, or preventing pregnancy in a mammal. They can also be used for

CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing

CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a

CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age

CC -related macular degeneration, multiple sclerosis and diabetes. The

CC products can also be used for detection and diagnosis and in bioreactors.

CC AAY60592 to AAY64572 represent specifically claimed peptides, and

CC AAY64573 to AAY64643 and AA233183 to AA233186 represent sequences used in

CC the exemplification of the present invention

XX

SO Sequence 7 AA:

Query Match 77.3%; Score 34; DB 3; Length 7;

Best Local Similarity 71.4%; Pred. No. 2e+06;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10

DB 1 CDPXTGC 7

RESULT 8

AAV62007

ID AAY62007 standard; peptide; 7 AA.

XX

AC AAY62007;

XX

DT 02-MAR-2000 (first entry)

XX

DE Cadherin-12 cell adhesion recognition cyclic peptide SEQ ID NO:1799.

XX

KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;

KW inhibition; cadherin extracellular domain; cell adhesion recognition;

KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;

KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;

KW cadherin related neuronal receptor; LI-cadherin; protocadherin;

KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;

KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;

KW neurological disease; cyclic.

XX

OS Synthetic.

OS Homo sapiens.

XX

FT Key Location/Qualifiers

FT Disulfide-bond 1..7

XX

PN WO9957149-A2.

```

XX 11-NOV-1999.
PD
XX
XX 05-MAY-1999; 99WO-CA000363.
PF
XX 05-MAY-1998; 98US-00073040.
PR
XX 06-NOV-1998; 98US-00187859.
PR
XX 20-JAN-1999; 99US-00234395.
PR
XX 08-MAR-1999; 99US-00264516.
XX
XX (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
XX Blaschuk OW, Gour BJ, Byers S;
XX
XX WPI; 2000-038791/03.
XX
XX New cadherin modulating agents, used for modulating nonclassical cadherin
XX mediated functions for treating e.g. cancers, obesity, rheumatoid
XX arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
XX Claim 46; Page 180; 252pp; English.
XX
XX The present invention describes cadherin modulating agents (MA)
XX comprising peptides which comprise a nonclassical cadherin cell adhesion
XX recognition (CAR) sequence. The MAs can be used for modulating
XX nonclassical cadherin-mediated functions. They can be used for e.g.
XX inhibiting adhesion of nonclassical-cadherin expressing cells in a
XX mammal, enhancing delivery of a drug through the skin of a mammal,
XX enhancing delivery of a drug to a tumour in a mammal, treating cancer in
XX a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
XX angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
XX expressing cell, preventing or treating obesity in a mammal, stimulating
XX blood vessel regression in a mammal, enhancing drug delivery to the
XX central nervous system, treating a demyelinating neurological disease,
XX increasing vasopermeability in a mammal, enhancing adhesion of
XX nonclassical cadherin-expressing cells, inhibiting synaptic stability in
XX a mammal, or preventing pregnancy in a mammal. They can also be used for
XX e.g. enhancing or directing neurite outgrowth, facilitating wound healing
XX or reducing scar tissue, or enhancing adhesion of foreign tissue in a
XX mammal. They can also be used for treating e.g. psoriasis, arthritis, age
XX -related macular degeneration, multiple sclerosis and diabetes. The
XX products can also be used for detection and diagnosis and in bioreactors.
XX CC AAY60592 to AAY64572 represent specifically claimed peptides, and
XX CC AAY64573 to AAY64643 and AA233183 to AA233186 represent sequences used in
XX the exemplification of the present invention
XX
XX Sequence 7 AA;
XX
XX Query Match 77.3%; Score 34; DB 3; Length 7;
XX Best Local Similarity 71.4%; Pred. No. 2e+06;
XX Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 4 CXPXTGC 10
XX | | | | |
XX 1 CDPXTGC 7
XX
XX Db
XX
XX RESULT 9
XX AAY62224
XX ID AAY62224 standard; peptide; 7 AA.
XX
XX AC AAY62224;
XX
XX 02-MAR-2000 (first entry)
XX
XX Cadherin-14 cell adhesion recognition cyclic peptide SEQ ID NO:3971.
XX
XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;
XX inhibition; cadherin extracellular domain; cell adhesion recognition;
XX OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
XX cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
XX cadherin related neuronal receptor; Li-cadherin; protocadherin;
XX desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;

```

```

XX Rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
XX neurological disease; cyclic.
XX
XX Synthetic.
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Disulfide-bond 1..7
XX
XX WO9597149-A2.
XX
XX 11-NOV-1999.
XX
XX 05-MAY-1999; 99WO-CA000363.
XX
XX 05-MAY-1998; 98US-00073040.
XX
XX 06-NOV-1998; 98US-00187859.
XX
XX 20-JAN-1999; 99US-00234395.
XX
XX 08-MAR-1999; 99US-00264516.
XX
XX (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
XX Blaschuk OW, Gour BJ, Byers S;
XX
XX WPI; 2000-038791/03.
XX
XX New cadherin modulating agents, used for modulating nonclassical cadherin
XX mediated functions for treating e.g. cancers, obesity, rheumatoid
XX arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
XX Claim 54; Page 184; 252pp; English.
XX
XX The present invention describes cadherin modulating agents (MA)
XX comprising peptides which comprise a nonclassical cadherin cell adhesion
XX recognition (CAR) sequence. The MAs can be used for modulating
XX nonclassical cadherin-mediated functions. They can be used for e.g.
XX inhibiting adhesion of nonclassical-cadherin expressing cells in a
XX mammal, enhancing delivery of a drug through the skin of a mammal,
XX enhancing delivery of a drug to a tumour in a mammal, treating cancer in
XX a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
XX angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
XX expressing cell, preventing or treating obesity in a mammal, stimulating
XX blood vessel regression in a mammal, enhancing drug delivery to the
XX central nervous system, treating a demyelinating neurological disease,
XX increasing vasopermeability in a mammal, enhancing adhesion of
XX nonclassical cadherin-expressing cells, inhibiting synaptic stability in
XX a mammal, or preventing pregnancy in a mammal. They can also be used for
XX e.g. enhancing or directing neurite outgrowth, facilitating wound healing
XX or reducing scar tissue, or enhancing adhesion of foreign tissue in a
XX mammal. They can also be used for treating e.g. psoriasis, arthritis, age
XX -related macular degeneration, multiple sclerosis and diabetes. The
XX products can also be used for detection and diagnosis and in bioreactors.
XX CC AAY60592 to AAY64572 represent specifically claimed peptides, and
XX CC AAY64573 to AAY64643 and AA233183 to AA233186 represent sequences used in
XX the exemplification of the present invention
XX
XX Sequence 7 AA;
XX
XX Query Match 77.3%; Score 34; DB 3; Length 7;
XX Best Local Similarity 71.4%; Pred. No. 2e+06;
XX Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 4 CXPXTGC 10
XX | | | | |
XX 1 CDPXTGC 7
XX
XX Db
XX
XX RESULT 10
XX ABU00550
XX ID ABU00550 standard; peptide; 7 AA.
XX
XX AC ABU00550;
XX

```

DT 05-SEP-2002 (first entry)  
XX  
DB B lymphocyte stimulator protein binding peptide #1.  
XX  
KW B lymphocyte stimulator protein binding protein; Blys; immune disease;  
KW allergy; proliferative disease; infectious disease; arteriosclerosis;  
KW inflammatory disorder; hypergammaglobulinaemia; blood clotting;  
KW ischaemia; graft-versus-host disease; neurodegenerative disease;  
KW immunosuppressive; nephrotropic; antirheumatic; antiarthritic;  
KW neuroprotective; cytostatic; immunostimulant; antitumour; anti-HIV;  
KW antiasclerotic; antiallergic; thyromimetic; antianaemic; haemostatic;  
KW dermatological; antihistaminic; cardiant; ophthalmological; uropathic;  
KW antidiabetic; antithyroid; antidepressant; hepatotropic.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 2 /label= Phe, Trp, Tyr  
FT FT  
FT Misc-difference 4 /label= Pro, Tyr  
FT FT  
XX  
PN WO200216411-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 17-AUG-2001; 2001WO-US025850.  
XX  
PR 18-AUG-2000; 2000US-0226700P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Belzer JP, Potter DM, Fleming TJ, Rosen CA;  
XX  
DR MPI; 2002-499775/53.  
XX  
PT The treatment of various diseases e.g. rheumatoid arthritis, comprises  
XX administering B lymphocyte stimulator binding polypeptide.  
XX  
PS Claim 69; Page 233; 387pp; English.  
XX  
XX The present invention relates to the treatment, prevention or  
CC amelioration of a disease or disorder associated with: aberrant B  
CC lymphocyte stimulator (Blys), Blys receptor expression or activity; cells  
CC of haematopoietic origin; or proliferative disease; and reducing,  
CC inhibiting or stimulating immunoglobulin production, B cell proliferation  
CC and graft rejection involving administration of Blys binding polypeptide.  
CC The Blys binding polypeptides are used in the treatment, prevention or  
CC amelioration of diseases such as immune system diseases, proliferative  
CC diseases, diseases of cells of haematopoietic origin, graft rejection,  
CC allergies, infectious diseases, arteriosclerosis, inflammatory disorders,  
CC hypergammaglobulinaemia, blood clotting disorders, ischaemia, and  
CC neurodegenerative diseases. The present sequence is a B lymphocyte  
CC stimulator protein binding peptide  
XX  
SQ Sequence 7 AA:  
Query Match 77.3%; Score 34; DB 5; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 CXPXTGC 10  
DB 1 CXPXTGC 7  
RESULT 11  
ABG33862  
ID ABG33862 standard; peptide; 7 AA.  
XX  
AC ABG33862;  
XX  
DT 15-JUL-2002 (first entry)

XX  
DE B lymphocyte stimulator (Blys) binding peptide #436.  
XX  
DB B lymphocyte stimulator protein; B lymphocyte stimulator binding peptide;  
KW Blys; biological fluid; serum; plasma; lymph; blood; urine; spinal fluid;  
KW synovial fluid; saliva; mucus.  
XX  
OS Synthetic.  
XX  
PN WO200216412-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 17-AUG-2001; 2001WO-US025891.  
XX  
PR 18-AUG-2000; 2000US-0226489P.  
XX  
PA (DYAX-) DYAX CORP.  
XX  
PI Belzer JP, Potter MD, Fleming TJ, Iadner RC;  
XX  
DR MPI; 2002-351647/38.  
XX  
PT New B lymphocyte stimulator binding polypeptide useful in detecting or  
XX isolating Blys or Blys-like polypeptide comprises a specified amino acid  
XX sequence.  
XX  
PS Disclosure; Page 132; 263pp; English.  
XX  
XX The invention relates to a B lymphocyte stimulator (Blys) binding  
CC polypeptide. Blys binding peptides bind Blys or Blys-like proteins  
CC reversibly or irreversibly. The binding peptides are used in detection,  
CC isolation and/or purification of Blys in a solution such as water or a  
CC buffer solution, as well as any fluid and/or cell obtained from an  
CC individual biological fluid, body tissue, body cell, cell line, tissue  
CC culture or other source containing Blys or Blys-like polypeptides. The  
CC biological fluids include sera, plasma, lymph, blood, blood fraction,  
CC urine, synovial fluid, spinal fluid, saliva and mucous. Sequences  
CC ABG33406-33415, ABG33423-33575, ABG33588-33846, ABG33848-33850 and  
CC ABG33852-33862 represent Blys binding peptides of the invention  
XX  
SQ Sequence 7 AA:  
Query Match 77.3%; Score 34; DB 5; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2e+06;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 CXPXTGC 10  
DB 1 CXPXTGC 7  
RESULT 12  
ABJ00545  
ID ABJ00545 standard; peptide; 13 AA.  
XX  
AC ABJ00545;  
XX  
DT 05-SEP-2002 (first entry)  
XX  
DB B lymphocyte stimulator protein binding protein consensus peptide #1.  
XX  
XX B lymphocyte stimulator protein binding protein; Blys; immune disease;  
KW allergy; proliferative disease; infectious disease; arteriosclerosis;  
KW inflammatory disorder; hypergammaglobulinaemia; blood clotting;  
KW ischaemia; graft-versus-host disease; neurodegenerative disease;  
KW immunosuppressive; nephrotropic; antirheumatic; antiarthritic;  
KW neuroprotective; cytostatic; immunostimulant; antitumour; anti-HIV;  
KW antiasclerotic; antiallergic; thyromimetic; antianaemic; haemostatic;  
KW dermatological; antihistaminic; cardiant; ophthalmological; uropathic;  
KW antidiabetic; antithyroid; antidepressant; hepatotropic.  
XX  
OS Unidentified.

```
XX Key Location/Qualifiers
FH Misc-difference 1 /label= Ala, Asn, Lys, Ser
FT Misc-difference 2 /label= Ala, Glu, Met, Ser, Val
FT Misc-difference 3 /label= Ala, Asn, Lys, Pro
FT Misc-difference 5 /label= Phe, Trp, Tyr
FT Misc-difference 7 /label= Pro, Tyr
FT Misc-difference 11 /label= Ala, Gln, His, Phe, Val
FT Misc-difference 12 /label= Asn, Gln, Gly, His, Ser, Val
FT Misc-difference 13 /label= Ala, Asn, Gly, Ile, Pro, Ser
FT WO200216411-A2.
XX 28-FEB-2002.
XX 17-AUG-2001; 2001WO-US025850.
XX 18-AUG-2000; 2000US-0226700P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Belzer JP, Potter DM, Fleming TL, Rosen CA;
XX WPI; 2002-499775/53.
XX The treatment of various diseases e.g. rheumatoid arthritis, comprises
PT administering B lymphocyte stimulator binding polypeptide.
XX Claim 69; Page 230; 387pp; English.
XX The present invention relates to the treatment, prevention or
CC amelioration of a disease or disorder associated with: aberrant B
CC lymphocyte stimulator (Blys), Blys receptor expression or activity; cells
CC of haematopoietic origin; or proliferative disease; and reducing,
CC inhibiting or stimulating immunoglobulin production, B cell proliferation
CC and graft rejection involving administration of Blys binding polypeptide.
CC The Blys binding polypeptides are used in the treatment, prevention or
CC amelioration of diseases such as immune system diseases, proliferative
CC diseases, diseases of cells of hematopoietic origin, graft rejection,
CC allergies, infectious diseases, arteriosclerosis, inflammatory disorders,
CC hypergammaglobulinaemia, blood clotting disorders, ischaemia, and
CC neurodegenerative diseases. The present sequence is a conserved region of
CC a B lymphocyte stimulator protein binding peptide
XX SQ
XX Sequence 13 AA;
XX Query Match 77.3%; Score 34; DB 5; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 23;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CXPXTGC 10
DB 4 CXPXTGC 10
RESULT 13
ABG33861
ID ABG33861 standard; peptide; 13 AA.
XX
XX ABG33861;
AC
XX 15-JUL-2002 (first entry)
DT
XX B lymphocyte stimulator (Blys) binding peptide #435.
DE
XX
```

```
KW B lymphocyte stimulator protein; B lymphocyte stimulator binding peptide;
KW Blys; biological fluid; serum; plasma; lymph; blood; urine; spinal fluid;
KW synovial fluid; saliva; mucus.
XX
OS Synthetic.
XX WO200216412-A2.
XX 28-FEB-2002.
XX 17-AUG-2001; 2001WO-US025891.
XX 18-AUG-2000; 2000US-0226489P.
XX (DVAX-) DVAX CORP.
XX Belzer JP, Potter MD, Fleming TL, Ladner RC;
XX WPI; 2002-351647/38.
XX
XX New B-lymphocyte stimulator binding polypeptide useful in detecting or
PT isolating Blys or Blys-like polypeptide comprises a specified amino acid
PT sequence.
XX Disclosure; Page 121; 269pp; English.
XX
XX The invention relates to a B lymphocyte stimulator (Blys) binding
CC polypeptide. Blys binding peptides bind Blys or Blys-like proteins
CC reversibly or irreversibly. The binding peptides are used in detection,
CC isolation and/or purification of Blys in a solution such as water or a
CC buffer solution, as well as any fluid and/or cell obtained from an
CC individual biological fluid, body tissue, body cell, cell line, tissue
CC culture or other source containing Blys or Blys-like polypeptides. The
CC biological fluids include sera, plasma, lymph, blood, blood fraction,
CC urine, synovial fluid, spinal fluid, saliva and mucus. Sequences
CC ABG33406-33415, ABG33423-33575, ABG33588-33846, ABG33848-33850 and
CC ABG33852-33862 represent Blys binding peptides of the invention
XX SQ
XX Sequence 13 AA;
XX Query Match 77.3%; Score 34; DB 5; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 23;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 CXPXTGC 10
DB 4 CXPXTGC 10
RESULT 14
AAU44932
ID AAU44932 standard; protein; 53 AA.
XX
XX AAU44932;
AC
XX 27-FEB-2002 (first entry)
DT
XX Propionibacterium acnes immunogenic protein #5828.
DE
XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
XX WO200181581-A2.
XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-US012865.
XX 21-APR-2000; 2000US-0199047P.
XX
```



XX US2004034888-A1.  
 PN  
 XX  
 XX 19-FEB-2004.  
 PD  
 XX  
 XX 28-APR-2003; 2003US-00425114.  
 PF  
 XX 06-MAY-1999; 99US-00304517.  
 PR 05-NOV-2001; 2001US-00985678.  
 XX  
 XX (LITUJ/) LITU J.  
 PA (ZHOV/) ZHOV Y.  
 PA (KOVA/) KOVALIC D K.  
 PA (SCRE/) SCREEN S E.  
 PA (TABAS/) TABASKA J E.  
 PA (CAOY/) CAO Y.  
 XX  
 XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;  
 PI  
 XX WPI; 2004-180133/17.  
 DR  
 XX  
 XX New recombinant DNA construct, useful for improving plant tolerance to  
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
 PT pests, for conferring increased resistance to plant disease, or for  
 PT improving yield.  
 PT  
 XX  
 XX Claim 1; SEQ ID NO 50232; 15pp; English.  
 PS  
 XX The invention describes a recombinant DNA construct comprising a  
 CC polynucleotide consisting of a sequence encoding an amino acid sequence  
 CC available in electronic form from the US patent office at  
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide  
 CC of the invention are also useful in physical arrays of molecules and as  
 CC plant breeding markers. The recombinant DNA construct is useful for  
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme  
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in  
 CC plant cells by modification of the cell cycle pathway, for conferring  
 CC increased resistance to plant disease, for producing galactomannan,  
 CC lignin or plant growth regulators, for increasing the rate of homologous  
 CC recombination in plants, for improving yield by modification of  
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake  
 CC or by providing improved plant growth and development under at least one  
 CC stress condition or for modifying seed oil or protein yield and/or  
 CC content. This is the amino acid sequence of a plant full length insert  
 CC polypeptide that can be used in the recombinant DNA construct of the  
 CC invention.  
 CC  
 XX  
 XX Sequence 99 AA;  
 SQ  
 Query Match 77.3%; Score 34; DB 8; Length 99;  
 Best Local Similarity 71.4%; Pred. No. 1.2e+02; Mismatches 2; Gaps 0;  
 Matches 5; Conservative 0; Indels 0;  
 OY 4 CXPXTGC 10  
 | | | | |  
 Db 91 CXPXTGC 97  
 RESULT 17  
 ADA54961  
 ID ADA54961 standard; protein; 113 AA.  
 XX  
 XX ADA54961;  
 AC  
 XX  
 XX 20-NOV-2003 (first entry)  
 DT  
 XX  
 XX Human protein, SEQ ID 2529.  
 DE  
 XX  
 XX Cytostatic; Anti-inflammatory; Osteopachic; Neuroprotective; Nootropic;  
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
 KW inflammatory disease; osteoporosis; neurological disease.  
 XX  
 XX Homo sapiens.  
 OS

XX  
 PN EP1293569-A2.  
 XX  
 XX 19-MAR-2003.  
 PD  
 XX  
 XX 21-MAR-2002; 2002EP-00006586.  
 PF  
 XX 14-SEP-2001; 2001JP-00328381.  
 PR 24-JAN-2002; 2002US-0350435P.  
 XX  
 XX (HELI-) HELIX RES INST.  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,  
 PI Yamamoto Y, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Maehno Y;  
 XX  
 XX WPI; 2003-395539/38.  
 DR  
 XX N-PSDB; ADA53322.  
 DR  
 XX  
 XX New polynucleotides encoding full-length polypeptides, e.g. secretory  
 PT and/or membrane proteins, useful for developing medicines for diseases in  
 PT which the gene is involved, or as target molecules for gene therapy.  
 PT  
 XX  
 XX Claim 14; SEQ ID NO 2529; 205pp; English.  
 PS  
 XX The present invention relates to novel human secretory or membrane  
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.  
 CC  
 XX  
 XX Sequence 113 AA;  
 SQ  
 Query Match 77.3%; Score 34; DB 6; Length 113;  
 Best Local Similarity 71.4%; Pred. No. 1.4e+02; Mismatches 2; Gaps 0;  
 Matches 5; Conservative 0; Indels 0;  
 OY 4 CXPXTGC 10  
 | | | | |  
 Db 103 CXPXTGC 109  
 RESULT 18  
 AAB42027  
 ID AAB42027 standard; protein; 138 AA.  
 XX  
 XX AAB42027;  
 AC  
 XX  
 XX 08-FEB-2001 (first entry)  
 DT  
 XX  
 XX Human ORF1791 polypeptide sequence SEQ ID NO:1582.  
 DE  
 XX  
 XX Human; open reading frame; ORF; detection; cytostatic; hepatotropic;  
 KW vulnary; antipapillary; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopachic; antiarthritis; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypoid;  
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200058473-A2.  
 PN  
 XX 05-OCT-2000.  
 PD

XX 31-MAR-2000; 2000WO-US008621.  
 PF 31-MAR-1999; 99US-0127607P.  
 XX 02-APR-1999; 99US-0127636P.  
 PR 05-APR-1999; 99US-0127728P.  
 PR 30-MAR-2000; 2000US-00540763.  
 XX (CURA-) CURAGEN CORP.  
 XX  
 PI Shinkets RA, Leach M;  
 DR WPI; 2000-602362/57.  
 DR N-PSDB; AAC76236.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease.  
 XX  
 PS Claim 11; Page 2743; 5507BP; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
 CC antiproliferative; antiparkinsonian; neuroprotective; osteopathic;  
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;  
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
 CC antiviral; antifungal; antineumatic; antithyroid; and antianemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 SQ Sequence 138 AA;  
 Query Match 77.3%; Score 34; DB 3; Length 138;  
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 CXPXTGC 10  
 Db 15 CGPRTGC 21  
 RESULT 19  
 ABO75329  
 ID ABO75329 standard; protein; 250 AA.  
 XX ABO75329;  
 AC  
 XX 29-JUL-2004 (first entry)  
 DT  
 XX Pseudomonas aeruginosa polypeptide #7504.  
 DE  
 XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
 KM  
 XX Pseudomonas aeruginosa.  
 OS  
 XX US6551795-B1.  
 XX  
 PN 22-APR-2003.  
 XX  
 PD 18-FEB-1999; 99US-00252991.  
 PF  
 XX

PR 18-FEB-1998; 98US-0074788P.  
 PR 27-JUL-1998; 98US-0094190P.  
 XX  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 PA Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
 XX  
 PI  
 XX  
 DR WPI; 2003-615309/58.  
 DR N-PSDB; ABD08900.  
 XX  
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
 PT pathological conditions resulting from bacterial infection.  
 XX  
 PS Disclosure; SEQ ID NO 24075; 455bp; English.  
 XX  
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful in diagnosis and  
 CC therapy of pathological conditions, as molecular targets for diagnostics,  
 CC prophylaxis and treatment of pathological conditions resulting from a  
 CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-P. aeruginosa drugs, as templates for recombinant  
 CC production of P. aeruginosa-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of P. aeruginosa-caused  
 CC infection, and in detection of P. aeruginosa sequences or other sequences  
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
 CC ABO68396 represent P. aeruginosa polypeptides of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html  
 XX  
 SQ Sequence 250 AA;  
 Query Match 77.3%; Score 34; DB 7; Length 250;  
 Best Local Similarity 71.4%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 CXPXTGC 10  
 Db 27 CGPRTGC 33  
 RESULT 20  
 ABB93198  
 ID ABB93198 standard; protein; 312 AA.  
 XX ABB93198;  
 AC  
 XX 31-MAY-2002 (first entry)  
 DT  
 XX Herbicidally active polypeptide SEQ ID NO 2409.  
 DE  
 XX Herbicidal; plant; agriculture; herbicide.  
 KM  
 XX Arabidopsis thaliana.  
 OS  
 XX WO200210210-A2.  
 XX  
 PN 07-FEB-2002.  
 PD  
 XX 28-AUG-2001; 2001WO-EP009892.  
 PF  
 XX 28-AUG-2001; 2001WO-EP009892.  
 PR  
 XX (FARB ) BAYER AG.  
 XX  
 PA Tiejien K, Weidler M;  
 XX  
 PI WPI; 2002-269010/31.  
 DR  
 XX Identifying plant target proteins for herbicidally active compounds,  
 PT

PT comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
XX organisms.  
XX  
PS Claim 5; SEQ ID NO 2409; 261dp + Sequence Listing; English.  
XX  
CC The invention relates to identifying target proteins (AB90790-AB94016)  
CC for herbicidally active compounds, comprising aligning and comparing  
CC nucleic acid or amino acid sequences from plant with nucleic acid or  
CC amino acid sequences from non-plant organisms using suitable search  
CC parameters, where plant sequences having an E-value greater by a factor  
CC of 3 than the E-value of most similar non-plant sequences are selected.  
CC The polypeptides or nucleic acids encoding them are useful for  
CC identifying modulators. The identified modulators are useful as  
CC herbicides  
XX  
SQ Sequence 312 AA;  
XX  
Query Match 77.3%; Score 34; DB 5; Length 312;  
Best Local Similarity 71.4%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 4 CXPXTGC 10  
DB 208 CDPNTGC 214  
XX  
RESULT 21  
ADN17632  
ID ADN17632 standard; protein; 532 AA.  
XX  
AC ADN17632;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polypeptide #285.  
XX  
KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
KM Bacteria.  
XX  
OS  
XX  
PN US2003233675-A1.  
XX  
PD 18-DEC-2003.  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
PR 21-FEB-2002; 2002US-0360039P.  
XX  
PA (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX  
DR WPI; 2004-061375/06.  
XX  
PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.  
XX  
PS Claim 1; SEQ ID NO 285; 122dp; English.  
XX  
CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
SQ Sequence 532 AA;  
XX  
Query Match 77.3%; Score 34; DB 8; Length 532;  
Best Local Similarity 71.4%; Pred. No. 4.8e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 4 CXPXTGC 10  
DB 122 CXPVTGC 128  
XX  
RESULT 22  
ADP05702  
ID ADP05702 standard; protein; 540 AA.  
XX  
AC ADP05702;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Bacterial polypeptide #1815.  
XX  
KM Proteus mirabilis infection; bacterial infection; antibacterial;  
KW immunostimulant.  
XX  
OS Proteus mirabilis.  
XX  
PN US6605709-B1.  
XX  
PD 12-AUG-2003.  
XX  
PF 05-APR-2000; 2000US-00543681.  
XX  
PR 09-APR-1999; 99US-0128706P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Breton GL;  
XX  
DR WPI; 2003-895291/82.  
XX  
PT New Proteus mirabilis polypeptides and polynucleotides, useful as  
PT reagents for diagnosis of bacterial disease, as components of  
PT antibacterial vaccines, as targets for antibacterial drugs, or as  
PT biocontrol agents for plants.  
XX  
PS Disclosure; SEQ ID NO 5987; 870dp; English.  
XX  
CC The invention relates to new Proteus mirabilis polypeptides and  
CC polynucleotides. The invention also relates to antibodies against the



CC polypeptides, methods for producing the polypeptides, a method of  
CC generating vaccines for immunising an individual against *P. mirabilis*, a  
CC method for evaluating a compound for the ability to bind a *P. mirabilis*  
CC polypeptide and a method for screening test compounds for anti-bacterial  
CC activity. The polypeptides and polynucleotides are useful as molecular  
CC targets for diagnosing, preventing and treating pathological conditions  
CC resulting from bacterial infection, as reagents for diagnosis of  
CC bacterial diseases, as components of antibacterial vaccines, as targets  
CC for antibacterial drugs or as bio-control agents for plants. This  
CC sequence represents a *Proteus mirabilis* polypeptide of the invention.

XX SQ Sequence 540 AA;

Query Match 77.3%; Score 34; DB 7; Length 540;  
Best Local Similarity 71.4%; Pred. No. 4.9e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
| | | | |  
Db 130 CXPVTGC 136

#### RESULT 23

ABG28408  
ID ABG28408 standard; protein; 3226 AA.

XX AC ABG28408;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #28399.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSB-) HYSBQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS92595.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
XX PT diagnostics, forensics, gene mapping, identification of mutations  
XX PT responsible for genetic disorders or other traits and to assess  
XX PT biodiversity.

XX PS Claim 20; SEQ ID NO 58767; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 3226 AA;

Query Match 77.3%; Score 34; DB 4; Length 3226;  
Best Local Similarity 71.4%; Pred. No. 2.1e+03;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
| | | | |  
Db 1707 CXPRTGC 1713

#### RESULT 24

AAU40006  
ID AAU40006 standard; protein; 158 AA.

XX AC AAU40006;

XX DT 13-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #902.

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US012865.

XX PR 21-APR-2000; 2000US-0199047P.

XX PR 02-JUN-2000; 2000US-0208841P.

XX PR 07-JUL-2000; 2000US-0216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX DR WPI; 2001-616774/71.

XX DR N-PSDB; AAS59510.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for  
XX PT vaccinating against and diagnosing infections, especially useful for  
XX PT treating acne vulgaris.

XX PS Example 1; SEQ ID NO 1201; 1069pp; English.

CC Sequences AAU9105-AAU6017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 158 AA;

Query Match 75.0%; Score 33; DB 4; Length 158;

Best Local Similarity 71.4%; Pred. No. 2.7e+02; Mismatches 2; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTC 10  
 DB 17 CXPXTC 23

RESULT 25

ID ABM36525 standard; protein; 158 AA.

XX ABM36525;

XX 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #1201.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;

XX immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

XX Zhang Y, Wang S, Jen S, Lodes WD, Benson DR, Jones R, Carter D;

XX Barch B, Valliave-Douglas J;

XX WPI, 2003-381789/36.

XX N-PSDB; ACF64439.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the  
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
 PT or for stimulating an immune response specific for a P. acnes protein.

XX Example 1; SEQ ID NO 1201; 1481bp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
 CC encoding a Propionibacterium acnes protein. The invention also relates to  
 CC polypeptides encoded by the polynucleotides (ABM36524-ABM64536) and to  
 CC immunogenic fragments of P. acnes polypeptides. The invention  
 CC additionally encompasses expression vectors and host cells comprising a  
 CC polynucleotide of the invention; antibodies against polypeptides of the  
 CC invention; fusion proteins comprising a polypeptide of the invention; a  
 CC method for stimulating an immune response specific for a P. acnes  
 CC polypeptide and an isolated T cell population comprising T cells prepared  
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,  
 CC polynucleotides, adjuvants, fusion proteins, T cell populations, or  
 CC antigen-presenting cells that express the polypeptide); a method and kit  
 CC for detecting or determining the presence or absence of P. acnes in a  
 CC patient; and a method for inhibiting the development of P. acnes in a  
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

CC proteins, T cell populations or antigen-presenting cells that express the  
 CC polypeptides are useful for diagnosing, preventing or treating acne  
 CC vulgaris, or for stimulating an immune response specific for a P. acnes  
 CC protein. The polynucleotides can also be used as probes or primers for  
 CC nucleic acid hybridisation. The vaccine composition is useful for the  
 CC stimulation of an immune response against P. acnes, or for treating acne,  
 CC and the kit is useful for performing a diagnostic assay. The present  
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
 CC reading frame) contained within the P. acnes polynucleotides of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 158 AA;

Query Match 75.0%; Score 33; DB 6; Length 158;

Best Local Similarity 71.4%; Pred. No. 2.7e+02; Mismatches 2; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTC 10  
 DB 17 CXPXTC 23

RESULT 26

ID ADT57626 standard; protein; 239 AA.

XX ADT57626;

XX 13-JAN-2005 (first entry)

XX Plant polypeptide, SEQ ID 7703.

XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;

XX disease resistance; galactomannan production; plant growth regulator;

XX heat tolerance; herbicide tolerance; lignin production;

XX extreme osmotic condition tolerance; pathogens resistance;

XX pest resistance; yield improvement; seed oil yield; seed protein yield.

XX Viridiplantae.

XX US2004216190-A1.

XX 28-OCT-2004.

XX 18-DEC-2003; 2003US-00739930.

XX 28-APR-2003; 2003US-00424599.

XX 28-APR-2003; 2003US-00425115.

XX (KOVA/) KOVALIC D K.

XX Kovalic DK;

XX WPI; 2004-757369/74.

XX New recombinant DNA constructs useful in the field of biochemistry and

PT genetics, and in particular for producing transgenic plants with improved

PT biological characteristics.

XX Claim 2; SEQ ID NO 7703; 14bp; English.

XX The invention relates a recombinant DNA construct comprising a  
 CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:  
 CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences  
 CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,  
 CC Arabidopsis, wheat and rape but the specification does not indicate which  
 CC sequences is derived from which organism. Also included is a method of  
 CC producing a plant having an improved property, comprising transforming a  
 CC plant with a recombinant DNA construct comprising a promoter region  
 CC functional in a plant cell operably joined to a polynucleotide encoding a  
 CC polypeptide associated with the property, and growing the transformed

CC plant. The property is selected from improving plant cold tolerance, for  
CC manipulating growth rate in plant cells by modification of the cell cycle  
CC pathway, for improving plant drought tolerance, for providing increased  
CC resistance to plant disease, for galactomannan production, for production  
CC of plant growth regulators, for improving plant heat tolerance, for  
CC improving plant tolerance to herbicides, for increasing the rate of  
CC homologous recombination in plants, for lignin production, for improving  
CC plant tolerance to extreme osmotic conditions, for improving plant  
CC tolerance to pathogens or pests, for yield improvement by modification of  
CC photosynthesis, for modifying seed oil yield and/or content, for  
CC modifying seed protein yield and/or content, for yield improvement by  
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake  
CC and for yield improvement by providing improved plant growth and  
CC development under at least one stress condition. The polynucleotide may  
CC also encode a plant transcription factor. The methods and compositions of  
CC the present invention are useful in the field of biochemistry and  
CC genetics, in particular for producing transgenic plants with improved  
CC biological characteristics such as increased yield, improved nitrogen  
CC flow, increasing plant tolerance to cold or heat, improving plant  
CC tolerance to extreme osmotic and drought conditions, and improving plant  
CC tolerance to plant pests or pathogens. They can also be used in physical  
CC arrays of molecules, plant breeding markers, computer-based storage and  
CC analysis systems. The present sequence is one of the 5544 plant protein  
CC sequences of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from USPTO at  
CC segdata.uspto.gov/sequence.html?docID=20040216190.

XX Sequence 239 AA:

Query Match 75.0%; Score 33; DB 8; Length 239;  
Best Local Similarity 71.4%; Pred. No. 3.8e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
| | | | |  
Db 141 CKPPTGC 147

RESULT 27

AAG43626 AAG43626 standard; protein; 751 AA.

XX AAG43626;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 54550.

KM Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301433.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144614P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149829P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157573P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160810P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 72.7%; Score 32; DB 3; Length 751;  
Best Local Similarity 57.1%; Pred. No. 1.5e+03;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTC 10  
| | : | |  
Db 90 CSPXTC 96

RESULT 28

HA643625  
ID AAG43625 strand: protein, 798 AA.

XX AAG43625;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 54549.

XX Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

OS Arabidopsis thaliana.

PN EP103405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153707P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.

PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160880P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161559P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Query Match 72.7%; Score 32; DB 3; Length 798;  
Best Local Similarity 57.1%; Pred. No. 1.6e+03;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 4 CXPXTCG 10  
| | | | |  
| | | | |  
Db 137 CSPTSGC 143  
  
RESULT 29  
ID AAG43624 standard; protein; 835 AA.  
XX AAG43624;  
XX  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54548.  
XX  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridization assay; genetic mapping; gene expression control; promoter;  
termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135362P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138647P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 18-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148655P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 23-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151067P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151338P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 25-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.

PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 29-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 72.7%; Score 32; DB 3; Length 835;  
 Best Local Similarity 57.1%; Pred. No. 1.6e+03;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CKPXTGC 10  
 Db 174 CSPTSGC 180

RESULT 30  
 AAY63264  
 ID AAY63264 standard; peptide; 7 AA.  
 AC AAY63264;  
 XX 02-MAR-2000 (first entry)  
 DT  
 XX  
 DE Protocadherin cell adhesion recognition cyclic peptide SEQ ID NO:2748.  
 XX  
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;  
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;  
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;  
 KW cadherin-14; cadherin-15; T-cadherin; PB-cadherin;  
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;  
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;  
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;  
 KW neurological disease; cyclic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FT Key Location/Qualifiers  
 FT Disulfide-bond 1..7  
 PN WO957149-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 PP 05-MAY-1999; 99WO-CA000363.  
 XX  
 PR 05-MAY-1998; 98US-00073040.  
 PR 06-NOV-1998; 98US-00187859.  
 PR 20-JAN-1999; 99US-00234395.  
 PR 08-MAR-1999; 99US-00264516.  
 XX  
 PA (ADBE-) ADHEREX TECHNOLOGIES INC.  
 XX  
 PI Blaschuk OM, Gour BJ, Byers S;  
 XX WPI, 2000-038791/03.  
 DR  
 XX  
 PT New cadherin modulating agents, used for modulating nonclassical cadherin  
 PT mediated functions for treating e.g. dancers, obesity, rheumatoid  
 PT arthritis, multiple sclerosis, diabetes or a neurological disease.  
 XX  
 PS Claim 84; Page 201; 252pp; English.  
 XX  
 CC The present invention describes cadherin modulating agents (MA)  
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion  
 CC recognition (CAR) sequence. The MAs can be used for modulating  
 CC nonclassical cadherin-mediated functions. They can be used for e.g.  
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a  
 CC mammal, enhancing delivery of a drug through the skin of a mammal,  
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in

CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting  
CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-  
CC expressing cell, preventing or treating obesity in a mammal, stimulating  
CC blood vessel regression in a mammal, enhancing drug delivery to the  
CC central nervous system, treating a demyelinating neurological disease,  
CC increasing vasopermeability in a mammal, enhancing synaptic stability in  
CC nonclassical cadherin-expressing cells, inhibiting pregnancy in a  
CC mammal, or preventing pregnancy in a mammal. They can also be used for  
CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing  
CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a  
CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age  
CC -related macular degeneration, multiple sclerosis and diabetes. The  
CC products can also be used for detection and diagnosis and in bioreactors.  
CC AA60592 to AA64572 represent specifically claimed peptides, and  
CC AA64573 to AA64643 and AA633183 to AA633186 represent sequences used in  
CC the exemplification of the present invention

SQ Sequence 7 AA;

Query Match 70.5%; Score 31; DB 3; Length 7;  
Best Local Similarity 57.1%; Pred. No. 2e+06;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
| | | |  
| : ||  
Db 1 CDPSSGC 7

RESULT 31

ABO81251  
ID ABO81251 standard; protein; 90 AA.

AC ABO81251;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #13426.

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

DR N-PSDB; ABD14822.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 29997; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target

CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84336 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the pinned  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html

SQ Sequence 90 AA;

Query Match 70.5%; Score 31; DB 7; Length 90;  
Best Local Similarity 57.1%; Pred. No. 3.9e+02;  
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
| | | |  
| : ||  
Db 29 CTPSAGC 35

RESULT 32

ABP06368  
ID ABP06368 standard; protein; 114 AA.

AC ABP06368;

DT 25-JUN-2002 (first entry)

DE Human ORFX protein sequence SEQ ID NO:12718.

KW Human, open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
KW hypertension; hypothyroidism; cholesterol ester storage disease;  
KW immune deficiency; immune disorder; infectious disease;  
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
KW myaesthesia gravis.

XX Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

PF 29-MAY-2001; 2001WO-US010836.

PR 30-MAY-2000; 2000US-0206132P.

PR 29-AUG-2000; 2000US-0228716P.

XX (CURA-) CURAGEN CORP.

PI Shinkets RA, Leach MD;

DR WPI; 2002-106308/14.

DR N-PSDB; ABR22120.

PT Novel human polypeptides and polynucleotides useful for diagnosing,  
PT preventing and treating cardiovascular disease, neurodegenerative,  
PT hyperproliferative disorders and autoimmune disorders.

XX Disclosure; SEQ ID NO 12718; 1037bp; English.

XX The present invention describes substantially purified human proteins  
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
XX in the specification). ABR15762 to ABR27252 encode the human ORFX  
XX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
XX treating or preventing a pathology associated with an ORFX-associated  
XX disorder in humans, and in the manufacture of a medicament for treating a  
XX syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
XX sequences can be used in gene therapy. ORFX sequences can be used in the  
XX treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
XX psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
XX osteoarthritis, neurodegenerative disorders, disorders related to organ



CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
 CC storage disease, various immune deficiencies and disorders, infectious  
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 CC disease and autoimmune inflammatory eye disease. OBPX proteins are also  
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 CC bone degenerative disorders, or periodontal disease, and for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues and conditions resulting from  
 CC systemic cytokine damage. N.B. The sequence data for this patent did not  
 CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 SQ Sequence 114 AA;

Query Match 70.5%; Score 31; DB 5; Length 114;  
 Best Local Similarity 57.1%; Pred. No. 4.7e+02;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
 | | | |  
 Db 83 CGPTSGC 89

## RESULT 33

ADJ71040 standard; protein; 148 AA.

ADJ71040;  
 06-MAY-2004 (first entry)

Human heat mitochondrial protein as a therapeutic target SegID2846.

XX mitochondrial; human; screening assay; diabetes mellitus;  
 XX Huntington's disease; osteoarthritis;  
 XX Leber's hereditary optic neuropathy; LHON;  
 XX mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
 XX myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
 XX neuroprotective; neurotropic; antidiabetic; anticonvulsant; antiarthritic;  
 XX osteopathic; ophthalmological; cytostatic.

OS Homo sapiens.

XX WO2003087768-A2.

XX 23-OCT-2003.

PF 04-APR-2003; 2003WO-US010870.

PR 12-APR-2002; 2002US-0372843P.

PR 17-JUN-2002; 2002US-038987P.

PR 20-SEP-2002; 2002US-0412418P.

XX (MINO-) MITOKOR.

PA (BUCK-) BUCK INST AGE RES.

PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;  
 PI Warnock DE;

DR WPI; 2003-845369/78.

PT Identifying a mitochondrial target for drug screening assays and for  
 PT treating diseases associated with altered mitochondrial function,  
 PT comprises detecting a modified polypeptide in a sample and correlating  
 PT with the disease.

PS Claim 1; SEQ ID NO 2846; 1800P; English.

CC This invention relates to novel mitochondrial targets that can be used  
 CC for therapeutic intervention in treating a disease associated with  
 CC altered mitochondrial function. Specifically, it refers to a method for

CC identifying proteins of the human heart mitochondrial proteome that are  
 CC useful for drug screening assays, as well as therapeutic targets. The  
 CC present invention describes a method for identifying such proteins that  
 CC can be used in the treatment of various diseases associated with altered  
 CC mitochondrial function including diabetes mellitus, Huntington's disease,  
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
 CC compositions have neuroprotective, neurotropic, antidiabetic,  
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
 CC cytosolic activities. This polypeptide sequence is a human heart  
 CC mitochondrial protein of the invention.

SQ Sequence 148 AA;

Query Match 70.5%; Score 31; DB 7; Length 148;  
 Best Local Similarity 57.1%; Pred. No. 5.9e+02;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
 | | | |  
 Db 69 CAPRAGC 75

## RESULT 34

ADP58925 standard; protein; 183 AA.

ADP58925;  
 12-FEB-2004 (first entry)

Human polypeptide sequence SEQ ID NO:1333.

XX biological activity; genetic engineering; hybridization probe; oligomer;  
 XX primer; chromosome mapping; gene mapping; recombinant protein production;  
 XX human.

OS Homo sapiens.

PN WO2003080795-A2.

PD 02-OCT-2003.

PF 09-AUG-2002; 2002WO-US025485.

PR 09-AUG-2001; 2001US-0311261P.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Yang Y, Wang Z, Weng G, Ma Y;

DR WPI; 2003-876918/81.

DR N-PSDB; ADP57925.

PT New polynucleotides, useful as hybridization probes, oligomers or  
 PT primers, for chromosome or gene mapping, for the recombinant production  
 PT of proteins, and for generating antisense DNA or RNA.

PS Claim 20; SEQ ID NO 1333; 571P; English.

CC The present sequence represents a polypeptide (II) with biological  
 CC activity, which is encoded by an isolated polynucleotide sequence (I)  
 CC from the present invention. Also described: (1) a vector comprising (I);  
 CC (2) an expression vector comprising (I); (3) a host cell genetically  
 CC engineered to comprise (I) which is operatively associated with a  
 CC regulatory sequence that modulates expression of (I) in the host cell;  
 CC (4) a polypeptide (II) encoded by (I); (5) a composition comprising the  
 CC polypeptide of (4) and a carrier; (6) an antibody directed against the  
 CC sample; (8) identifying a compound that binds to the polypeptide of (4) in a  
 CC sample; (9) producing the polypeptide of (4); and (10) a collection of  
 CC polynucleotides comprising at least one of the polynucleotide sequences

CC (1). The polynucleotides (I) can be used as hybridisation probes,  
 CC oligomers or primers, for chromosome or gene mapping, for the recombinant  
 CC production of proteins, and for generating antisense DNA or RNA.

XX  
 SQ Sequence 183 AA;

Query Match 70.5%; Score 31; DB 7; Length 183;  
 Best Local Similarity 57.1%; Pred. No. 7e+02;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10  
 | | : | |  
 Db 96 CSPLSGC 102

RESULT 35  
 ID AAO00015  
 AA000015 standard; protein; 193 AA.

XX AC AA000015;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 13907.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KM nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR N-PSDB; AA179946.

XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 XX and treating e.g. leukemia, inflammation and immune disorders.

XX PS Claim 20; SEQ ID NO 13907; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and  
 CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 193 AA;

Query Match 70.5%; Score 31; DB 4; Length 193;  
 Best Local Similarity 57.1%; Pred. No. 7.3e+02;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10  
 | | : | |  
 Db 98 CSPLSGC 104

RESULT 36  
 ID ADF60298  
 ADF60298 standard; protein; 193 AA.

XX AC ADF60298;

XX DT 12-FEB-2004 (first entry)

XX DE Human config polypeptide sequence SEQ ID NO:2665.

XX biological activity; genetic engineering; hybridisation probe; oligomer;  
 KM primer; chromosome mapping; gene mapping; recombinant protein production;  
 KM human.

XX OS Homo sapiens.

XX PN WO2003080795-A2.

XX PD 02-OCT-2003.

XX PF 09-AUG-2002; 2002WO-US025485.

XX PR 09-AUG-2001; 2001US-0311261P.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Yang Y, Wang Z, Weng G, Ma Y;

XX DR WPI: 2003-876918/81.

XX DR N-PSDB; ADF59846.

XX New polynucleotides, useful as hybridization probes, oligomers or  
 XX primers, for chromosome or gene mapping, for the recombinant production  
 XX of proteins, and for generating antisense DNA or RNA.  
 XX Example 3; SEQ ID NO 2665; 571pp; English.

XX The present invention describes isolated polynucleotide sequences (I),  
 CC which encode polypeptides (II) with biological activity. Also described:  
 CC (1) a vector comprising (I); (2) an expression vector comprising (I); (3)  
 CC a host cell genetically engineered to comprise (I) which is operatively  
 CC associated with a regulatory sequence that modulates expression of (I) in  
 CC the host cell; (4) a polypeptide (II) encoded by (I); (5) a composition  
 CC comprising the polypeptide of (4) and a carrier; (6) an antibody directed  
 CC against the polypeptide of (4); (7) detecting (I) or the polypeptide of  
 CC (4) in a sample; (8) identifying a compound that binds to the polypeptide  
 CC of (4); (9) producing the polypeptide of (4); and (10) a collection of  
 CC polynucleotides comprising at least one of the polynucleotide sequences  
 CC (I). The polynucleotides (I) can be used as hybridisation probes,  
 CC oligomers or primers, for chromosome or gene mapping, for the recombinant  
 CC production of proteins, and for generating antisense DNA or RNA. The  
 CC present sequence represents a human config polypeptide sequence, which is  
 CC used in an example from the present invention.

XX SQ Sequence 193 AA;

Query Match 70.5%; Score 31; DB 7; Length 193;  
 Best Local Similarity 57.1%; Pred. No. 7.3e+02;  
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 CXPXTGC 10  
 | | : | |  
 Db 98 CSPLSGC 104

RESULT 37  
 ID ABP99259  
 ABP99259 standard; protein; 238 AA.

```

XX AC ABP99259;
XX XX
XX DT 21-MAR-2003 (first entry)
XX DE Orthosomycin biosynthetic polypeptide SEQ ID NO 105.
XX KW Orthosomycin; biosynthesis; evernimycin; avilamycin; enzyme.
XX OS Streptomyces mobaraensis.
XX PN WO200279505-A2.
XX PD 10-OCT-2002.
XX PF 26-MAR-2002; 2002WO-CA000432.
XX PR 28-MAR-2001; 2001US-0279095P.
XX PR 30-MAR-2001; 2001US-0279709P.
XX PR 20-APR-2001; 2001US-0285214P.
XX PS (ECOP-) ECOPIA BIOSCIENCES INC.
XX PI Farnet CM, Zazopoulos E, Staffa A;
XX DR WPI; 2003-058435/05.
XX DR N-PSDB; ABZ66722.
XX PT Identifying orthosomycin biosynthetic gene, gene fragment or gene
XX PT cluster, by detecting presence of nucleic acid sequence corresponding to
XX PT 17 of flamamycins protein families.
XX PS Claim 1; Page 260-261, 511pp; English.
XX CC The invention relates to identifying orthosomycin biosynthetic genes and
XX CC its fragment/gene cluster (ABZ6670-ABZ66813), comprising detecting the
XX CC presence of a nucleic acid sequence coding for a polypeptide (ABP99207-
XX CC ABP99362). The method is useful for identifying an orthosomycin
XX CC biosynthetic gene, gene fragment or gene cluster, especially an
XX CC evernimycin-type or avilamycin-type orthosomycin biosynthetic gene,
XX CC gene fragment or gene cluster. The method is useful for detecting the
XX CC presence of any organism that contains DNA for the production of
XX CC orthosomycins (both evernimycin-type orthosomycins and avilamycin-type
XX CC orthosomycins) regardless of the level at which genes for orthosomycin
XX CC production are expressed by the organism or the amount of orthosomycin
XX CC produced by the organism. This allows for the detection of new
XX CC orthosomycin natural products, not produced by the organism
XX CC
XX SQ Sequence 238 AA;
XX
XX Query Match 70.5%; Score 31; DB 6; Length 238;
XX Best Local Similarity 57.1%; Pred. No. 8.7e+02;
XX Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 4 CXPXTGC 10
XX | | : | |
XX 26 CAPDSCC 32
XX
XX RESULT 38
XX ABP99348
XX ID ABP99348 standard; protein; 239 AA.
XX XX
XX AC ABP99348;
XX XX
XX DT 21-MAR-2003 (first entry)
XX DE Orthosomycin biosynthetic polypeptide GenBank No AAK83184.
XX KW Orthosomycin; biosynthesis; evernimycin; avilamycin; enzyme.
XX OS Streptomyces viridochromogenes.
XX

```

```

XX PN WO200279505-A2.
XX XX
XX PD 10-OCT-2002.
XX XX
XX PF 26-MAR-2002; 2002WO-CA000432.
XX PR 28-MAR-2001; 2001US-0279095P.
XX PR 30-MAR-2001; 2001US-0279709P.
XX PR 20-APR-2001; 2001US-0285214P.
XX PS (ECOP-) ECOPIA BIOSCIENCES INC.
XX PI Farnet CM, Zazopoulos E, Staffa A;
XX DR WPI; 2003-058435/05.
XX PT Identifying orthosomycin biosynthetic gene, gene fragment or gene
XX PT cluster, by detecting presence of nucleic acid sequence corresponding to
XX PT 17 of flamamycins protein families.
XX PS Claim 1; Page; 511pp; English.
XX CC The invention relates to identifying orthosomycin biosynthetic genes and
XX CC its fragment/gene cluster (ABZ6670-ABZ66813), comprising detecting the
XX CC presence of a nucleic acid sequence coding for a polypeptide (ABP99207-
XX CC ABP99362). The method is useful for identifying an orthosomycin
XX CC biosynthetic gene, gene fragment or gene cluster, especially an
XX CC evernimycin-type or avilamycin-type orthosomycin biosynthetic gene,
XX CC gene fragment or gene cluster. The method is useful for detecting the
XX CC presence of any organism that contains DNA for the production of
XX CC orthosomycins (both evernimycin-type orthosomycins and avilamycin-type
XX CC orthosomycins) regardless of the level at which genes for orthosomycin
XX CC production are expressed by the organism or the amount of orthosomycin
XX CC produced by the organism. This allows for the detection of new
XX CC orthosomycin natural products, not produced by the organism. Note: The
XX CC present sequence is not given in the specification but was obtained from
XX CC GenBank
XX SQ Sequence 239 AA;
XX
XX Query Match 70.5%; Score 31; DB 6; Length 239;
XX Best Local Similarity 57.1%; Pred. No. 8.7e+02;
XX Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 4 CXPXTGC 10
XX | | : | |
XX 22 CAPDSCC 28
XX
XX RESULT 39
XX ABP76713
XX ID ABP76713 standard; protein; 239 AA.
XX XX
XX AC ABP76713;
XX XX
XX DT 26-FEB-2003 (first entry)
XX DE Streptomyces viridochromogenes Avig6.
XX KW Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;
XX KW medicine; Staphylococcus aureus; biosynthetic gene cluster; enzyme.
XX OS Streptomyces viridochromogenes.
XX PN WO200268436-A1.
XX PD 06-SEP-2002.
XX PF 24-AUG-2001; 2001WO-BP009815.
XX PR 25-FEB-2001; 2001DE-01009166.
XX PS (COMB-) COMBINATURE BIOPHARM AG.
XX

```

XX  
PI Weinauer G, Muehlenweg A, Treifzer A, Bechtold A;  
XX WPI; 2003-018650/01.  
DR N-PSDB; AB237547.  
XX  
PT New avilamycin derivatives, useful for treatment of infections, and  
PT nucleic acid encoding avilamycin synthesis enzymes.  
PS Claim 16; Page 187-189; 31pp; German.  
XX  
CC The invention relates to avilamycin derivatives (I) with antibacterial,  
CC virucide, protozoacide and fungicide activity. (I) are useful for  
CC treatment of infections (bacterial, viral, protozoal or fungal), in human  
CC or veterinary medicine, particularly where caused by *Staphylococcus*  
CC aureus. (I) are more hydrophilic than known avilamycins. The present  
CC sequence is that of an avilamycin synthesis enzyme from the *Streptomyces*  
CC viridochromogenes Avilamycin A biosynthetic gene cluster (AB237515-  
CC AB237516)  
CC  
SQ Sequence 239 AA;  
  
Query Match 70.5%; Score 31; DB 6; Length 239;  
Best Local Similarity 57.1%; Pred. No. 8.7e+02;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 4 CXPXTGC 10  
| | : ||  
DB 22 CAPDSCG 28  
  
RESULT 40  
ADX95230  
ID ADX95230 standard; protein; 285 AA.  
XX  
AC ADX95230;  
XX  
DT 21-APR-2005 (first entry)  
XX  
DE Plant full length insert polypeptide seqid 57894.  
XX  
XX plant protectant; plant growth regulant; gene therapy; plant;  
XX recombinant DNA construct; physical array; plant breeding marker;  
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
XX extreme osmotic condition; pathogen tolerance; pest tolerance;  
XX growth rate; cell cycle pathway; disease resistance;  
XX galactomannan production; lignin production; plant growth regulator;  
XX yield; plant growth; plant development; seed oil; protein yield;  
XX protein content.  
XX  
OS Unidentified.  
XX  
XX US2004034888-A1.  
XX PN  
XX 19-FEB-2004.  
XX PD  
XX 28-APR-2003; 2003US-00425114.  
XX PF  
XX 06-MAY-1999; 99US-00304517.  
XX PR  
XX 05-NOV-2001; 2001US-00985678.  
XX PR  
XX (LIU/) LIU J.  
XX PA (ZHOU/) ZHOU Y.  
XX PA (KOVA/) KOVALIC D K.  
XX PA (SCRE/) SCREEN S E.  
XX PA (TABAS/) TABASKA J E.  
XX PA (CAOV/) CAO Y.  
XX  
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;  
XX WPI; 2004-180133/17.  
XX  
PT New recombinant DNA construct, useful for improving plant tolerance to

PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
PT pests, for conferring increased resistance to plant disease, or for  
PT improving yield.  
PS Claim 1; SEQ ID NO 57894; 15pp; English.  
XX  
CC The invention describes a recombinant DNA construct comprising a  
CC polynucleotide consisting of a sequence encoding an amino acid sequence  
CC available in electronic form from the US patent office at  
CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide  
CC of the invention are also useful in physical arrays of molecules and as  
CC plant breeding markers. The recombinant DNA construct is useful for  
CC improving plant tolerance to cold, heat, drought, herbicides, extreme  
CC osmotic conditions, pathogens or pests, for manipulating growth rate in  
CC plant cells by modification of the cell cycle pathway, for conferring  
CC increased resistance to plant disease, for producing galactomannan,  
CC lignin or plant growth regulators, for increasing the rate of homologous  
CC recombination in plants, for improving yield by modification of  
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake  
CC or by providing improved plant growth and development under at least one  
CC stress condition or for modifying seed oil or protein yield and/or  
CC content. This is the amino acid sequence of a plant full length insert  
CC polypeptide that can be used in the recombinant DNA construct of the  
CC invention.  
CC  
SQ Sequence 285 AA;  
  
Query Match 70.5%; Score 31; DB 8; Length 285;  
Best Local Similarity 57.1%; Pred. No. 1e+03;  
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 4 CXPXTGC 10  
| | : ||  
DB 160 CAPLSCG 166  
  
RESULT 41  
ABO74638  
ID ABO74638 standard; protein; 310 AA.  
XX  
XX ABO74638;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Pseudomonas aeruginosa polypeptide #6813.  
XX  
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX Pseudomonas aeruginosa.  
XX  
XX US651795-B1.  
XX PN  
XX 22-APR-2003.  
XX PD  
XX 18-FEB-1999; 99US-00252991.  
XX PF  
XX 18-FEB-1998; 98US-0074788P.  
XX PR  
XX 27-JUL-1998; 98US-0094190P.  
XX PR  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX PA  
XX  
PI Rubenfield MJ, Nolling J, DeLoughery C, Bush D;  
XX WPI; 2003-615309/58.  
XX DR N-PSDB; ABD08209.  
XX  
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
XX disclosure; SEQ ID NO 23384; 45pp; English.  
XX  
CC The invention relates to Pseudomonas aeruginosa polypeptides and the

CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biotech technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html

XX SQ Sequence 310 AA;

XX Query Match 70.5%; Score 31; DB 7; Length 310;  
XX Best Local Similarity 57.1%; Pred. No. 1.1e+03;  
XX Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

XX QY 4 CXPXTGC 10  
XX | | | |  
XX 236 CSPAGC 242

XX DB

XX RESULT 42  
XX ADU17203  
XX ID ADU17203 standard; protein; 354 AA.

XX AC ADU17203;  
XX DT 13-JAN-2005 (first entry)

XX DE M. tuberculosis exported polypeptides #611.

XX XX M. tuberculosis exported polypeptides #611.

XX KW antitubercular; tuberculostatic; vaccine; tuberculosis; hepatitis.

XX OS Mycobacterium tuberculosis.

XX PN US2004214165-A1.

XX XX 28-OCT-2004.

XX PD 16-MAY-2001; 2001US-00855604.

XX PF 16-MAY-2001; 2001US-00855604.

XX PR 16-MAY-2001; 2001US-00855604.

XX XX (GICO/) GICOUET B.  
XX PA (PORT/) PORTNOI D.  
XX PA (LIME/) LIM E.  
XX PA (PELI/) PELICIC V.  
XX PA (GUIGU/) GUIGUENO A.  
XX PA (GOGU/) GOGUET D L S Y.

XX PI Gicouet B, Portnoi D, Lim E, Pelicic V, Guigueno A, Goguett DLSY;  
XX DR WPI; 2004-765598/75.  
XX DR N-PSDB; ADU17188.

XX PT Novel recombinant screening, cloning or expression vector replicating in  
XX PT mycobacteria, useful for preventing Mycobacterium tuberculosis, having  
XX PT replicon that is functional in mycobacteria, selectable marker and  
XX PT reporter cassette.

XX PS Example 5; SEQ ID NO 863; 220pp; English.

XX CC The invention relates to a recombinant screening, cloning or expression  
XX CC vector, that replicates in mycobacteria, comprising a replicon that is  
XX CC functional in mycobacteria, selectable marker, reporter cassette. The  
XX CC recombinant vector is useful for diagnosing and preventing M.  
XX CC tuberculosis. Also disclosed is a vaccine composition for tuberculosis

CC and hepatitis. The present sequence represents the amino acid sequence of  
CC a Mycobacterium tuberculosis exported polypeptide.

XX SQ Sequence 354 AA;

XX Query Match 70.5%; Score 31; DB 8; Length 354;  
XX Best Local Similarity 57.1%; Pred. No. 1.2e+03;  
XX Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

XX QY 4 CXPXTGC 10  
XX | | | |  
XX 80 CSPAGC 86

XX DB

XX RESULT 43  
XX ABO70819  
XX ID ABO70819 standard; protein; 444 AA.

XX AC ABO70819;  
XX XX 29-JUL-2004 (first entry)

XX DT Pseudomonas aeruginosa polypeptide #2994.

XX DE Pseudomonas aeruginosa polypeptide #2994.

XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX OS Pseudomonas aeruginosa.

XX PN US6551795-B1.

XX PD 22-APR-2003.

XX PF 18-FEB-1999; 99US-00252991.

XX PR 18-FEB-1998; 98US-0074788P.

XX PR 27-JUL-1998; 98US-0094190P.

XX XX (GENO-) GENOME THERAPEUTICS CORP.

XX PA Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX PI WPI; 2003-615309/58.  
XX DR N-PSDB; ABD04390.

XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX PT pathological conditions resulting from bacterial infection.

XX PS Disclosure; SEQ ID NO 19565; 455pp; English.

XX XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and  
XX CC therapy of pathological conditions, as molecular targets for diagnostics,  
XX CC prophylaxis and treatment of pathological conditions resulting from a  
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,  
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX CC effective antibacterial targets, as targets for antibacterial drugs,  
XX CC including anti-P. aeruginosa drugs, as templates for recombinant  
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target  
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused  
XX CC infection, and in detection of P. aeruginosa sequences or other sequences  
XX CC of Pseudomonas species using biotech technology. Sequences ABO67826-  
XX CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX CC sequence data for this patent did not form part of the printed  
XX CC specification but was obtained in electronic format from USPTO at  
XX CC seqdata.uspto.gov/sequence.html

XX SQ Sequence 444 AA;

XX Query Match 70.5%; Score 31; DB 7; Length 444;  
XX Best Local Similarity 57.1%; Pred. No. 1.5e+03;  
XX Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 CXPXTGC 10  
 XX | | | |  
 DB 47 CSPAAAGC 53

## RESULT 44

ABB93340  
 ID ABB93340 standard; protein; 482 AA.

XX ABB93340;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 2551.

XX Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP009892.

PR 28-AUG-2001; 2001WO-EP009892.

PA (FARB ) BAYER AG.

PI Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,

PT comprising aligning and comparing nucleic acid or amino acid sequences

PT from plant with nucleic acid or amino acid sequences from non-plant

XX organisms.

PS Claim 5; SEQ ID NO 2551; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins (ABB90790-ABB94016)

CC for herbicidally active compounds, comprising aligning and comparing

CC nucleic acid or amino acid sequences from plant with nucleic acid or

CC amino acid sequences from non-plant organisms using suitable search

CC parameters, where plant sequences having an E-value greater by a factor

CC of 3 than the E-value of most similar non-plant sequences are selected.

CC The polypeptides or nucleic acids encoding them are useful for

CC identifying modulators. The identified modulators are useful as

XX herbicides

XX Sequence 482 AA;

XX Query Match 70.5%; Score 31; DB 5; Length 482;

XX Best Local Similarity 57.1%; Pred. No. 1.6e+03;

XX Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 CXPXTGC 10

DB 301 CTPGSGC 307

XX RESULT 45

AAW32363

ID AAW32363 standard; protein; 580 AA.

XX AAW32363;

XX 13-JAN-1998 (first entry)

DE Mycobacterium tuberculosis antigen TBRa29.

XX Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;

KW skin testing; M.tuberculosis.

XX Mycobacterium tuberculosis.  
 OS  
 XX WO9709429-A2.  
 PN  
 XX 13-MAR-1997.  
 PD  
 XX 30-AUG-1996; 96WO-US014675.  
 PF  
 XX 01-SEP-1995; 95US-00523435.  
 PR 22-SEP-1995; 95US-00532136.  
 PR 22-MAR-1996; 96US-00620280.  
 PR 05-JUN-1996; 96US-00658800.  
 PR 12-JUL-1996; 96US-00680573.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 XX Reed SG, Skeiky YAM, Dillon DC, Campos-Neto A, Houghton R;  
 PI Vedvick TH, Twardzik DR;  
 XX  
 DR WPI: 1997-192904/17.  
 DR N-PSDB; AAT91410.  
 XX  
 XX New immunogenic polypeptide(s) from soluble M. tuberculosis antigens -  
 PT useful for diagnosis of M. tuberculosis infection.  
 XX  
 XX Example 3; Page 118-121, 190pp; English.

XX A new immunogenic polypeptide has been developed comprising an  
 CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its  
 CC variant differing only in conservative substitutions and/or  
 CC modifications). The present sequence represents a M.tuberculosis antigen,  
 CC TBRa29. The immunogenic polypeptide can be used to diagnose  
 CC M.tuberculosis infection by forming complexes with specific antibodies in  
 CC the sample. Fragments of DNA encoding the immunogenic polypeptide can be  
 CC used as diagnostic primers or probes and agents that bind to the antigen,  
 CC especially monoclonal antibodies or equivalent polyclonal antibodies, are  
 CC also used for diagnosis

XX Sequence 580 AA;

XX Query Match 70.5%; Score 31; DB 2; Length 580;

XX Best Local Similarity 57.1%; Pred. No. 1.8e+03;

XX Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 4 CXPXTGC 10

DB 58 CSPAVGC 64

XX RESULT 46

AAW32431

ID AAW32431 standard; protein; 580 AA.

XX AAW32431;

XX 08-JAN-1998 (first entry)

DE Mycobacterium tuberculosis antigen TBRa29.

XX Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;

KW skin testing; M.tuberculosis.

XX Mycobacterium tuberculosis.

XX WO9709428-A2.

XX 13-MAR-1997.

XX 30-AUG-1996; 96WO-US014674.

XX 01-SEP-1995; 95US-00523436.

PR 22-SEP-1995; 95US-00533634.

PR 22-MAR-1996; 96US-00620874.  
 PR 05-JUN-1996; 96US-00659683.  
 PR 12-JUL-1996; 96US-00680574.  
 XX  
 PA (CORI-) CORIXA CORP.  
 PI Reed SG, Skeiky YA, Dillon DC, Campos-Neto A, Houghton R;  
 PI Vedvick TS, Twardzik DR;  
 XX  
 DR WPI: 1997-192903/17.  
 DR N-PSDB; AAT91473.  
 XX  
 PT New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are  
 PT useful in vaccines for prevention or treatment of tuberculosis, also for  
 PT diagnosis.  
 XX  
 PS Example 3; Page 109-111; 168pp; English.  
 XX  
 CC A new immunogenic polypeptide has been developed comprising an  
 CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its  
 CC variant differing only in conservative substitutions and/or  
 CC modifications). The present sequence represents a M.tuberculosis antigen,  
 CC TBRa29. The immunogenic protein, and fusion proteins containing one or  
 CC more of the proteins or one of the proteins plus ESAT-6, are useful in  
 CC vaccines, preferably when formulated with a non-specific adjuvant, to  
 CC induce an immune response against M.tuberculosis (for treatment or  
 CC prevention)  
 CC  
 SQ Sequence 580 AA;

Query Match 70.5%; Score 31; DB 2; Length 580;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
 DB 58 CSPAVGC 64

RESULT 47  
 ID AAM64303  
 ID AAM64303 standard; protein; 580 AA.  
 AC AAM64303;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 09-NOV-1998 (first entry)  
 DB Mycobacterium tuberculosis antigen TBRa29.  
 XX  
 KW Tuberculosis; infection; diagnosis; antigen; TBRa29.  
 OS Mycobacterium tuberculosis; strain H37Ra.  
 PN WO9816645-A2.  
 XX  
 PD 23-APR-1998.  
 PF 07-OCT-1997; 97WO-US018214.  
 XX  
 PR 11-OCT-1996; 96US-00729622.  
 PR 13-MAR-1997; 97US-00818111.  
 XX  
 PA (CORI-) CORIXA CORP.  
 PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;  
 PI Vedvick TS, Twardzik DR, Lodes MJ;  
 XX  
 DR WPI: 1998-251292/22.  
 DR N-PSDB; AAV44351.  
 XX  
 PT New isolated Mycobacterium tuberculosis polypeptides and DNA - used to  
 PT develop products for the detection of M. tuberculosis infection and

PT diagnosis of tuberculosis.  
 XX  
 PS Example 3; Page 111-113; 250pp; English.  
 XX

CC This polypeptide comprises Mycobacterium tuberculosis soluble antigen  
 CC TBRa29. It is encoded by a DNA sequence (see AAV44351) isolated from a M.  
 CC tuberculosis strain H37Ra expression library with rabbit anti-sera raised  
 CC against M. tuberculosis supernatant. TBRa29 shows some homology to a  
 CC sequence previously identified in Mycobacterium leprae sequence but not  
 CC in M. tuberculosis. The invention relates to compositions and methods for  
 CC diagnosing tuberculosis. It provides polypeptides (see AAM64291-W64379)  
 CC comprising an antigenic portion of a soluble M. tuberculosis antigen, or  
 CC an immunogenic portion of an M. tuberculosis antigen, as well as DNA  
 CC sequences encoding such polypeptides, recombinant expression vectors and  
 CC transformed or transfected host cells. Also claimed are methods and  
 CC diagnostic kits for detecting M. tuberculosis infection in a patient  
 CC using the above polypeptides, antibodies or oligonucleotide probes and  
 CC primers, for the diagnosis of tuberculosis. (Updated on 17-OCT-2003 to  
 CC standardise OS field)

Query Match 70.5%; Score 31; DB 2; Length 580;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
 DB 58 CSPAVGC 64

RESULT 48  
 ID AAM81666  
 ID AAM81666 standard; protein; 580 AA.

AC AAM81666;  
 XX  
 DT 27-JAN-1999 (first entry)

DE M. tuberculosis immunogenic polypeptide TBRa29.

KW Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;  
 KW vaccine; pharmaceutical; infection; diagnosis.

OS Mycobacterium tuberculosis.

PN WO9816646-A2.

PD 23-APR-1998.

PF 07-OCT-1997; 97WO-US018293.

XX 11-OCT-1996; 96US-00730510.

PR 13-MAR-1997; 97US-00818112.

XX (CORI-) CORIXA CORP.

PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;  
 PI Vedvick TS, Twardzik DR, Lodes MJ;

XX WPI: 1998-261042/23.

DR N-PSDB; AAV64459.

XX

PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to  
 PT develop products for the detection of M. tuberculosis infection and for  
 PT diagnosis, treatment and prevention of tuberculosis.

XX Example 3; Page 107-108; 230pp; English.

CC This sequence represents an immunogenic portion of a soluble  
 CC Mycobacterium tuberculosis (MT) antigen which can be used in a method for  
 CC inducing protective immunity against tuberculosis (TB). This sequence can  
 CC be formulated into vaccines and/or pharmaceutical compositions for

CC immunising against M. tuberculosis infection or may be used for the  
 CC diagnosis of tuberculosis  
 XX  
 SQ Sequence 580 AA;

Query Match 70.5%; Score 31; DB 2; Length 580;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
 | | | |  
 DB 58 CSPAVGC 64

RESULT 49  
 AAY38968  
 ID AAY38968 standard; protein; 580 AA.

XX AAY38968;

XX 05-NOV-1999 (first entry)

XX M. tuberculosis recombinant antigen protein TBra29.

XX Antigen; diagnosis; detection; infection; antibody; immunisation;  
 KW vaccine; immunity.

OS Mycobacterium tuberculosis.

PN WO9942118-A2.

XX 26-AUG-1999.

XX 17-FEB-1999; 99WO-US003265.

XX 18-FEB-1998; 98US-00024753.

XX 05-MAY-1998; 98US-00072596.

XX (CORI-) CORIXA CORP.

XX Reed SG, Skeiky YAM, Dillon DC, Campos-Neto A, Houghton R;  
 PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;

XX WPI: 1999-527416/44.

XX N-PSDB; AA219049.

XX New polypeptide comprising antigenic portions of M. tuberculosis.

XX Example 3; Page 147-149; 323pp; English.

CC This invention describes novel recombinant antigens and their encoding  
 CC nucleic acids derived from Mycobacterium tuberculosis. The novel  
 CC polypeptides are useful for detecting M. tuberculosis infection in a  
 CC biological sample by detecting antibodies which bind with the  
 CC polypeptides, and are useful as vaccines for immunizing against M.  
 CC tuberculosis infection. The new detection methods are needed as current  
 CC vaccination strategies do not provide 100% immunity

XX Sequence 580 AA;

Query Match 70.5%; Score 31; DB 2; Length 580;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
 | | | |  
 DB 58 CSPAVGC 64

RESULT 50  
 AAY39105  
 ID AAY39105 standard; protein; 580 AA.

AC AAY39105;  
 XX  
 DT 05-NOV-1999 (first entry)

XX M. tuberculosis antigen TBra29 amino acid sequence.

XX Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;  
 KW immunotherapy; diagnosis; immunisation; vaccine; infection;  
 KW immune response; skin test.

OS Mycobacterium tuberculosis.

PN WO9942076-A2.

XX 26-AUG-1999.

XX 17-FEB-1999; 99WO-US003268.

XX 18-FEB-1998; 98US-00025197.

XX 05-MAY-1998; 98US-00072967.

XX (CORI-) CORIXA CORP.

XX Reed SG, Skeiky YAM, Dillon DC, Campos-Neto A, Houghton R;  
 PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;

XX WPI: 1999-527409/44.

XX N-PSDB; AA219261.

XX New antigens from Mycobacterium tuberculosis useful in diagnostic skin  
 PT tests and protective or therapeutic vaccines or compositions.

XX Example 3; Page 105-106; 299pp; English.

XX The present invention describes polypeptides comprising an immunogenic  
 CC part of a Mycobacterium tuberculosis antigen (Ag). Also described are  
 CC vaccines and fusion protein containing M. tuberculosis Ag's. M.  
 CC tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other  
 CC polypeptides fragments, can be used in pharmaceutical compositions or  
 CC vaccines to generate a protective or therapeutic immune response to M.  
 CC tuberculosis and as reagents in skin tests for diagnosis of tuberculosis.  
 CC Ag can induce proliferation of, or cytokine secretion by, T, B or natural  
 CC killer cells and/or macrophages in tuberculosis-immune subjects. AA219249  
 CC to AA219460 and AAY39083 to AAY39225 are used in the exemplification of  
 CC the present invention

XX Sequence 580 AA;

Query Match 70.5%; Score 31; DB 2; Length 580;  
 Best Local Similarity 57.1%; Pred. No. 1.8e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CXPXTGC 10  
 | | | |  
 DB 58 CSPAVGC 64

Search completed: January 4, 2006, 15:55:15  
 Job time : 128.965 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 ; Search time 95.8087 Seconds

(without alignments)  
64.204 Million cell updates/sec

Title: US-09-932-322-2

Perfect score: 30  
Sequence: 1 XXXXXXXXXXXXX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database :

A\_Geneseq\_21:.\*  
1: geneseqp1980s:.\*  
2: geneseqp1990s:.\*  
3: geneseqp2000s:.\*  
4: geneseqp2001s:.\*  
5: geneseqp2002s:.\*  
6: geneseqp2003as:.\*  
7: geneseqp2003bs:.\*  
8: geneseqp2004s:.\*  
9: geneseqp2005s:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	60.0	9	8	ADQ13022 Hepatitis
2	18	60.0	9	8	ADQ13021 Hepatitis
3	18	60.0	11	2	AAW65943 Molecule
4	18	60.0	11	2	AAW65944 Molecule
5	18	60.0	15	9	ADV23688 HBV Immun
6	18	60.0	15	9	ADV23687 HBV Immun
7	18	60.0	24	4	ABR42273 Peptide #
8	18	60.0	24	4	AAW36081 Peptide #
9	18	60.0	24	4	AAW75973 Human bon
10	18	60.0	24	4	AAW63159 Human bon
11	18	60.0	24	4	ABG57702 Human liv
12	18	60.0	69	4	ABR68695 Drosophil
13	18	60.0	69	8	ADP30715 Human sec
14	18	60.0	69	8	ABR39058 Human sec
15	18	60.0	91	4	AAW32549 Peptide #
16	18	60.0	91	4	AAW72290 Human bon
17	18	60.0	91	4	AAW59710 Human bon
18	18	60.0	91	4	ABG53976 Human bra
19	18	60.0	91	5	ABG42105 Human liv
20	18	60.0	93	8	ADP30859 Human pep
21	18	60.0	99	7	ABO83167 Human sec
22	18	60.0	99	8	ADP30709 Human sec
23	18	60.0	107	7	ABO74878 Pseudomon
24	18	60.0	115	7	ABO73394 Pseudomon

25	18	60.0	122	7	ADP04777 Bacterial
26	18	60.0	135	8	ADP30703 Human sec
27	18	60.0	144	7	ABO75730 Pseudomon
28	18	60.0	144	8	ADP31453 Human sec
29	18	60.0	144	8	ADP30450 Plant ful
30	18	60.0	150	6	ABU45351 Protein e
31	18	60.0	153	3	AAW41043 Human ORF
32	18	60.0	153	5	ABP09251 Human ORF
33	18	60.0	154	7	ABO74421 Pseudomon
34	18	60.0	156	7	ABO77836 Pseudomon
35	18	60.0	159	7	ADP30761 Human sec
36	18	60.0	162	8	ADP30758 Human sec
37	18	60.0	162	8	ADP30756 Human sec
38	18	60.0	162	8	ADP30757 Human sec
39	18	60.0	162	8	ADP30813 Human sec
40	18	60.0	165	8	ADP31269 Human sec
41	18	60.0	168	8	ADP31099 Human sec
42	18	60.0	171	8	ADP30794 Human sec
43	18	60.0	171	8	ADP30793 Human sec
44	18	60.0	177	3	AAW10193 Arabidops
45	18	60.0	177	8	ADP30755 Human sec
46	18	60.0	177	8	ADP31682 Human sec
47	18	60.0	177	8	ADP36379 Plant pol
48	18	60.0	182	7	ABM66593 Rice abio
49	18	60.0	185	7	ABM69559 Rice abio
50	18	60.0	192	8	ADP31576 Human sec
51	18	60.0	194	7	ABO78166 Pseudomon
52	18	60.0	199	8	ADY13615 Plant ful
53	18	60.0	201	8	ADP30711 Human sec
54	18	60.0	204	3	AAW10192 Arabidops
55	18	60.0	210	8	ADP30986 Human sec
56	18	60.0	210	8	ADP30960 Human sec
57	18	60.0	219	8	ADP30593 Human sec
58	18	60.0	222	8	ADP31354 Human sec
59	18	60.0	228	8	ADP30764 Human sec
60	18	60.0	233	3	AAW74791 Neisseria
61	18	60.0	234	8	ADP31480 Human sec
62	18	60.0	234	8	ADP31481 Human sec
63	18	60.0	237	8	ADP30792 Human sec
64	18	60.0	239	7	ABO81291 Pseudomon
65	18	60.0	249	8	ADP30754 Human sec
66	18	60.0	249	8	ADP31443 Human sec
67	18	60.0	252	8	ADP31485 Human sec
68	18	60.0	258	7	ABO80689 Pseudomon
69	18	60.0	258	8	ADP30597 Human sec
70	18	60.0	267	8	ADP30822 Human sec
71	18	60.0	270	7	ABO68843 Pseudomon
72	18	60.0	276	8	ADP30568 Human sec
73	18	60.0	277	8	ADP30503 Human sec
74	18	60.0	279	8	ADP31107 Human sec
75	18	60.0	291	8	ADP31193 Human sec
76	18	60.0	294	8	ADP31045 Human sec
77	18	60.0	300	8	ADP30775 Human sec
78	18	60.0	309	8	ADP30862 Human sec
79	18	60.0	309	8	ADP30873 Human sec
80	18	60.0	310	7	ABO80230 Pseudomon
81	18	60.0	312	8	ADP31454 Human sec
82	18	60.0	318	7	ABM65364 Mouse pro
83	18	60.0	318	8	ADP30588 Human sec
84	18	60.0	320	8	ADP30824 Human sec
85	18	60.0	320	8	ADP31607 Human sec
86	18	60.0	320	8	ADP31649 Human sec
87	18	60.0	323	7	ABO74543 Pseudomon
88	18	60.0	327	7	ABM69959 Rice abio
89	18	60.0	330	4	ABBS9598 Drosophil
90	18	60.0	339	8	ADP30892 Human sec
91	18	60.0	339	8	ADP30702 Human sec
92	18	60.0	345	8	ADP31016 Human sec
93	18	60.0	345	8	ADP31683 Human sec
94	18	60.0	346	8	ADP31683 Human sec
95	18	60.0	348	8	ADP31676 Human sec
96	18	60.0	348	8	ADP31374 Human sec
97	18	60.0	353	8	ADP31558 Human sec

98	18	60.0	354	8	ADP30783	Adp30783	Human sec	171	18	60.0	558	8	ADP31255	Adp31255	Human sec
99	18	60.0	354	8	ADP30707	Adp30707	Human sec	172	18	60.0	564	8	ADP31194	Adp31194	Human sec
100	18	60.0	354	8	ADP30779	Adp30779	Human sec	173	18	60.0	574	8	ADP31133	Adp31133	Human sec
101	18	60.0	354	8	ADP30784	Adp30784	Human sec	174	18	60.0	576	8	ADP31248	Adp31248	Human sec
102	18	60.0	354	8	ADP30778	Adp30778	Human sec	175	18	60.0	583	2	ADH11358	Vertebrat	
103	18	60.0	355	4	ABM57739	Abm57739	Drosophil	176	18	60.0	588	8	ADP31680	Adp31680	Human sec
104	18	60.0	355	7	ABM87551	Abm87551	Rice abio	177	18	60.0	592	8	ADP30917	Adp30917	Human sec
105	18	60.0	355	9	ADZ76101	Adz76101	Fruit fly	178	18	60.0	598	8	ADP30895	Adp30895	Human sec
106	18	60.0	357	9	ADP31263	Adp31263	Human sec	179	18	60.0	604	8	ADP311789	Adp311789	Human the
107	18	60.0	357	8	ADP31267	Adp31267	Human sec	180	18	60.0	603	8	ADP31150	Adp31150	Human sec
108	18	60.0	360	8	ADP31439	Adp31439	Human sec	181	18	60.0	604	8	ADP30940	Adp30940	Human sec
109	18	60.0	366	8	ADP31106	Adp31106	Human sec	182	18	60.0	605	8	ADP30941	Adp30941	Human sec
110	18	60.0	366	8	ADP31670	Adp31670	Human sec	183	18	60.0	605	8	ADP30507	Adp30507	Human sec
111	18	60.0	367	5	ADP31670	Adp31670	Human sec	184	18	60.0	609	8	ADP31263	Adp31263	Human sec
112	18	60.0	369	7	ADP31694	Adp31694	Human sec	185	18	60.0	612	8	ADP31064	Adp31064	Human sec
113	18	60.0	370	8	ABM88560	Abm88560	Rice abio	186	18	60.0	615	8	ADP31132	Adp31132	Human sec
114	18	60.0	372	8	ADY09767	Ady09767	Plant full	187	18	60.0	615	8	ADP31361	Adp31361	Human sec
115	18	60.0	382	8	ADP31610	Adp31610	Human sec	188	18	60.0	618	8	ADP31360	Adp31360	Human sec
116	18	60.0	382	8	ADP31216	Adp31216	Human sec	189	18	60.0	617	8	ADP31657	Adp31657	Human sec
117	18	60.0	382	8	ADP31221	Adp31221	Human sec	190	18	60.0	618	8	ADP31554	Adp31554	Human sec
118	18	60.0	387	8	ADP31158	Adp31158	Human sec	191	18	60.0	621	8	ADP30896	Adp30896	Human sec
119	18	60.0	390	8	ADP31218	Adp31218	Human sec	192	18	60.0	639	8	ADP31521	Adp31521	Human sec
120	18	60.0	393	8	ADP31345	Adp31345	Human sec	193	18	60.0	639	8	ADP30861	Adp30861	Human sec
121	18	60.0	399	8	ADP31464	Adp31464	Human sec	194	18	60.0	642	8	ADP31265	Adp31265	Human sec
122	18	60.0	401	7	ABO68344	AbO68344	Pseudomon	195	18	60.0	645	8	ADP31124	Adp31124	Human sec
123	18	60.0	411	8	ADP31104	Adp31104	Human sec	196	18	60.0	645	8	ADP31125	Adp31125	Human sec
124	18	60.0	420	8	ADP30729	Adp30729	Human sec	197	18	60.0	645	8	ADP30858	Adp30858	Human sec
125	18	60.0	421	8	ADP31349	Adp31349	Human sec	198	18	60.0	645	8	ADP31183	Adp31183	Human sec
126	18	60.0	423	8	ADP31479	Adp31479	Human sec	199	18	60.0	654	8	ADP30666	Adp30666	Human sec
127	18	60.0	425	8	ADP30549	Adp30549	Human sec	200	18	60.0	658	8	ADP31226	Adp31226	Human sec
128	18	60.0	429	8	ADP31207	Adp31207	Human sec	201	18	60.0	665	8	ADP30571	Adp30571	Human sec
129	18	60.0	434	5	ABP41972	Abp41972	Human ova	202	18	60.0	666	8	ADP31547	Adp31547	Human sec
130	18	60.0	441	8	ADP31206	Adp31206	Human sec	203	18	60.0	666	8	ADP31256	Adp31256	Human sec
131	18	60.0	449	8	ADQ21283	Adq21283	Human sof	204	18	60.0	669	8	ADP31598	Adp31598	Human sec
132	18	60.0	450	8	ADP30887	Adp30887	Human sec	205	18	60.0	669	8	ADP31493	Adp31493	Human sec
133	18	60.0	453	8	ADP31465	Adp31465	Human sec	206	18	60.0	677	8	ADP31585	Adp31585	Human sec
134	18	60.0	471	8	ADP31567	Adp31567	Human sec	207	18	60.0	681	8	ADP30868	Adp30868	Human sec
135	18	60.0	478	8	ADP31007	Adp31007	Human sec	208	18	60.0	681	8	ADP31053	Adp31053	Human sec
136	18	60.0	483	8	ADP30852	Adp30852	Human sec	209	18	60.0	699	8	ADP31080	Adp31080	Human sec
137	18	60.0	483	8	ADP30853	Adp30853	Human sec	210	18	60.0	699	8	ADP31079	Adp31079	Human sec
138	18	60.0	492	4	AA871707	Aa871707	Human ERR	211	18	60.0	699	8	ADP31241	Adp31241	Human sec
139	18	60.0	492	7	ADP314170	Adp314170	Human SRC	212	18	60.0	699	8	ADP31240	Adp31240	Human sec
140	18	60.0	494	6	ABR64215	AbR64215	Angiogene	213	18	60.0	702	8	ADP31518	Adp31518	Human sec
141	18	60.0	494	8	ADP319178	Ado19178	Human PRO	214	18	60.0	711	8	ADP31215	Adp31215	Human sec
142	18	60.0	494	8	ADP319174	Ado19174	Human PRO	215	18	60.0	711	8	ADP31652	Adp31652	Human sec
143	18	60.0	494	8	ADP54221	Adp54221	Human PRO	216	18	60.0	711	8	ADP31535	Adp31535	Human sec
144	18	60.0	494	9	ADP54221	Adp54221	Cyclin-de	217	18	60.0	714	8	ADP31440	Adp31440	Human sec
145	18	60.0	494	9	ADP54221	Adp54221	Cyclin-de	218	18	60.0	714	8	ADP31561	Adp31561	Human sec
146	18	60.0	494	9	ADP54221	Adp54221	Cyclin-de	219	18	60.0	720	5	ADP31494	Adp31494	Human sec
147	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	220	18	60.0	720	8	ADP31494	Adp31494	Human sec
148	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	221	18	60.0	722	8	ADP31494	Adp31494	Human sec
149	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	222	18	60.0	722	8	ADP31494	Adp31494	Human sec
150	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	223	18	60.0	722	8	ADP31494	Adp31494	Human sec
151	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	224	18	60.0	722	8	ADP31494	Adp31494	Human sec
152	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	225	18	60.0	722	8	ADP31494	Adp31494	Human sec
153	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	226	18	60.0	722	8	ADP31494	Adp31494	Human sec
154	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	227	18	60.0	722	8	ADP31494	Adp31494	Human sec
155	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	228	18	60.0	722	8	ADP31494	Adp31494	Human sec
156	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	229	18	60.0	722	8	ADP31494	Adp31494	Human sec
157	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	230	18	60.0	722	8	ADP31494	Adp31494	Human sec
158	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	231	18	60.0	722	8	ADP31494	Adp31494	Human sec
159	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	232	18	60.0	722	8	ADP31494	Adp31494	Human sec
160	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	233	18	60.0	722	8	ADP31494	Adp31494	Human sec
161	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	234	18	60.0	722	8	ADP31494	Adp31494	Human sec
162	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	235	18	60.0	722	8	ADP31494	Adp31494	Human sec
163	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	236	18	60.0	722	8	ADP31494	Adp31494	Human sec
164	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	237	18	60.0	722	8	ADP31494	Adp31494	Human sec
165	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	238	18	60.0	722	8	ADP31494	Adp31494	Human sec
166	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	239	18	60.0	722	8	ADP31494	Adp31494	Human sec
167	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	240	18	60.0	722	8	ADP31494	Adp31494	Human sec
168	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	241	18	60.0	722	8	ADP31494	Adp31494	Human sec
169	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	242	18	60.0	722	8	ADP31494	Adp31494	Human sec
170	18	60.0	502	7	ADP314502	Adp314502	PRO polyP	243	18	60.0	722	8	ADP31494	Adp31494	Human sec

244	18	60.0	807	8	ADP31036	Adp31036	Human	sec	317	18	60.0	1092	8	ADP31153	Adp31153	Human	sec
245	18	60.0	810	8	ADP31307	Adp31307	Human	sec	318	18	60.0	1092	8	ADP31358	Adp31358	Human	sec
246	18	60.0	813	8	ADP30649	Adp30649	Human	sec	319	18	60.0	1107	8	ADP30482	Adp30482	Human	sec
247	18	60.0	813	8	ADP30561	Adp30561	Human	sec	320	18	60.0	1107	8	ADP30504	Adp30504	Human	sec
248	18	60.0	816	8	ADP31249	Adp31249	Human	sec	321	18	60.0	1110	8	ADP31430	Adp31430	Human	sec
249	18	60.0	816	8	ADP31581	Adp31581	Human	sec	322	18	60.0	1113	8	ADP31508	Adp31508	Human	sec
250	18	60.0	828	8	ADP31569	Adp31569	Human	sec	323	18	60.0	1116	8	ADP31203	Adp31203	Human	sec
251	18	60.0	831	8	ADP31333	Adp31333	Human	sec	324	18	60.0	1116	8	ADP31692	Adp31692	Human	sec
252	18	60.0	831	8	ADP31123	Adp31123	Human	sec	325	18	60.0	1116	8	ADP31202	Adp31202	Human	sec
253	18	60.0	831	8	ADP31179	Adp31179	Human	sec	326	18	60.0	1128	6	ADA15725	AdA15725	C. elegans	
254	18	60.0	843	3	AA144348	Aay44348	Human	hep	327	18	60.0	1133	8	ADP30931	Adp30931	Human	sec
255	18	60.0	843	3	ADP30663	Adp30663	Human	sec	328	18	60.0	1133	8	ADP30928	Adp30928	Human	sec
256	18	60.0	849	8	ADP31182	Adp31182	Human	sec	329	18	60.0	1133	8	ADP30933	Adp30933	Human	sec
257	18	60.0	861	8	ADP31020	Adp31020	Human	sec	330	18	60.0	1133	8	ADP30939	Adp30939	Human	sec
258	18	60.0	861	8	ADP31021	Adp31021	Human	sec	331	18	60.0	1134	8	ADP30647	Adp30647	Human	sec
259	18	60.0	865	8	ADP30566	Adp30566	Human	sec	332	18	60.0	1140	8	ADP31128	Adp31128	Human	sec
260	18	60.0	869	7	AB068000	Ab068000	Pseudomon		333	18	60.0	1140	8	ADP31130	Adp31130	Human	sec
261	18	60.0	870	8	ADP30646	Adp30646	Human	sec	334	18	60.0	1141	8	ADP30677	Adp30677	Human	sec
262	18	60.0	882	8	ADP31688	Adp31688	Human	sec	335	18	60.0	1142	8	ADP30929	Adp30929	Human	sec
263	18	60.0	885	8	ADP31198	Adp31198	Human	sec	336	18	60.0	1155	8	ADP31455	Adp31455	Human	sec
264	18	60.0	887	8	ADP30554	Adp30554	Human	sec	337	18	60.0	1165	8	ADP30879	Adp30879	Human	sec
265	18	60.0	887	8	ADP30548	Adp30548	Human	sec	338	18	60.0	1168	8	ADP31046	Adp31046	Human	sec
266	18	60.0	889	8	ADP31648	Adp31648	Human	sec	339	18	60.0	1170	8	ADP30919	Adp30919	Human	sec
267	18	60.0	890	8	ADP31059	Adp31059	Human	sec	340	18	60.0	1170	8	ADP30922	Adp30922	Human	sec
268	18	60.0	892	8	ADP31578	Adp31578	Human	sec	341	18	60.0	1173	8	ADP31155	Adp31155	Human	sec
269	18	60.0	897	8	ADP30914	Adp30914	Human	sec	342	18	60.0	1189	8	ADP31043	Adp31043	Human	sec
270	18	60.0	906	8	ADP31409	Adp31409	Human	sec	343	18	60.0	1192	8	ADP30964	Adp30964	Human	sec
271	18	60.0	918	8	ADP31459	Adp31459	Human	sec	344	18	60.0	1192	8	ADP31180	Adp31180	Human	sec
272	18	60.0	921	8	ADP31522	Adp31522	Human	sec	345	18	60.0	1197	8	ADP31034	Adp31034	Human	sec
273	18	60.0	925	5	AA014246	Aa014246	Human	pre	346	18	60.0	1199	8	ADP31044	Adp31044	Human	sec
274	18	60.0	930	8	ADP31444	Adp31444	Human	sec	347	18	60.0	1215	8	ADP30898	Adp30898	Human	sec
275	18	60.0	933	8	ADP31510	Adp31510	Human	sec	348	18	60.0	1215	8	ADP30900	Adp30900	Human	sec
276	18	60.0	936	8	ADP31597	Adp31597	Human	sec	349	18	60.0	1221	8	ADP31293	Adp31293	Human	sec
277	18	60.0	936	8	ADP31568	Adp31568	Human	sec	350	18	60.0	1224	8	ADP31426	Adp31426	Human	sec
278	18	60.0	939	8	ADP31542	Adp31542	Human	sec	351	18	60.0	1233	8	ADP30523	Adp30523	Human	sec
279	18	60.0	939	8	ADP31086	Adp31086	Human	sec	352	18	60.0	1239	8	ADP31297	Adp31297	Human	sec
280	18	60.0	939	8	ADP31541	Adp31541	Human	sec	353	18	60.0	1252	8	ADP30678	Adp30678	Human	sec
281	18	60.0	945	8	ADP31238	Adp31238	Human	sec	354	18	60.0	1260	8	ADP31533	Adp31533	Human	sec
282	18	60.0	945	8	ADP31237	Adp31237	Human	sec	355	18	60.0	1269	8	ADP31382	Adp31382	Human	sec
283	18	60.0	947	8	ADP30937	Adp30937	Human	sec	356	18	60.0	1269	8	ADP31381	Adp31381	Human	sec
284	18	60.0	948	8	ADP30586	Adp30586	Human	sec	357	18	60.0	1282	8	ADP31328	Adp31328	Human	sec
285	18	60.0	950	8	ADP31167	Adp31167	Human	sec	358	18	60.0	1288	8	ADP31114	Adp31114	Human	sec
286	18	60.0	954	8	ADP31015	Adp31015	Human	sec	359	18	60.0	1302	8	ADP30675	Adp30675	Human	sec
287	18	60.0	957	8	ADP31528	Adp31528	Human	sec	360	18	60.0	1309	8	ADP31695	Adp31695	Human	sec
288	18	60.0	960	8	ADP31471	Adp31471	Human	sec	361	18	60.0	1312	8	ADP30999	Adp30999	Human	sec
289	18	60.0	960	8	ADP31470	Adp31470	Human	sec	362	18	60.0	1358	8	ADP30995	Adp30995	Human	sec
290	18	60.0	966	8	ADP30745	Adp30745	Human	sec	363	18	60.0	1359	8	ADP31129	Adp31129	Human	sec
291	18	60.0	975	8	ADP30843	Adp30843	Human	sec	364	18	60.0	1362	8	ADP31181	Adp31181	Human	sec
292	18	60.0	990	8	ADP31553	Adp31553	Human	sec	365	18	60.0	1365	8	ADP31035	Adp31035	Human	sec
293	18	60.0	1002	8	ADG39639	Adg39639	Human	pan	366	18	60.0	1371	8	ADP31646	Adp31646	Human	sec
294	18	60.0	1002	8	ADP30866	Adp30866	Human	sec	367	18	60.0	1371	8	ADP30876	Adp30876	Human	sec
295	18	60.0	1010	8	ADP31296	Adp31296	Human	sec	368	18	60.0	1380	8	ADP31566	Adp31566	Human	sec
296	18	60.0	1017	6	AA026719	Aa026719	SR	prote1	369	18	60.0	1383	8	ADP31091	Adp31091	Human	sec
297	18	60.0	1017	6	ADP31268	Adp31268	Human	sec	370	18	60.0	1387	8	ADP30946	Adp30946	Human	sec
298	18	60.0	1023	8	ADP30994	Adp30994	Human	sec	371	18	60.0	1404	8	ADP31627	Adp31627	Human	sec
299	18	60.0	1030	8	ADP30913	Adp30913	Human	sec	372	18	60.0	1420	8	ADP30944	Adp30944	Human	sec
300	18	60.0	1033	8	ADP30984	Adp30984	Human	sec	373	18	60.0	1431	8	ADP31509	Adp31509	Human	sec
301	18	60.0	1035	8	ADP31552	Adp31552	Human	sec	374	18	60.0	1437	8	ADP31357	Adp31357	Human	sec
302	18	60.0	1038	8	ADP30860	Adp30860	Human	sec	375	18	60.0	1440	8	ADG34533	Adg34533	Glucocort	
303	18	60.0	1048	8	ADP31642	Adp31642	Human	sec	376	18	60.0	1454	8	ADP31177	Adp31177	Human	sec
304	18	60.0	1056	8	ADP31082	Adp31082	Human	sec	377	18	60.0	1456	8	ADP30923	Adp30923	Human	sec
305	18	60.0	1057	8	ADP31592	Adp31592	Human	sec	378	18	60.0	1464	8	ADP31040	Adp31040	Human	sec
306	18	60.0	1059	8	ADP31042	Adp31042	Human	sec	379	18	60.0	1464	8	ADP31437	Adp31437	Human	sec
307	18	60.0	1065	8	ADP30973	Adp30973	Human	sec	380	18	60.0	1470	8	ADP31390	Adp31390	Human	sec
308	18	60.0	1065	8	ADP31347	Adp31347	Human	sec	381	18	60.0	1480	8	ADP31611	Adp31611	Human	sec
309	18	60.0	1076	6	AB081145	Ab081145	Human	PRO	382	18	60.0	1482	8	ADP30557	Adp30557	Human	sec
310	18	60.0	1076	6	ABU66845	Aby66845	Human	PRO	383	18	60.0	1485	8	ADP31383	Adp31383	Human	sec
311	18	60.0	1081	2	AA124319	Aay44319	Mouse	dep	384	18	60.0	1485	8	ADP31384	Adp31384	Human	sec
312	18	60.0	1083	8	ADP31073	Adp31073	Human	sec	385	18	60.0	1488	8	ADP31385	Adp31385	Human	sec
313	18	60.0	1086	8	ADP31175	Adp31175	Human	sec	386	18	60.0	1488	8	ADP31386	Adp31386	Human	sec
314	18	60.0	1086	8	ADP31447	Adp31447	Human	sec	387	18	60.0	1498	8	ADP30684	Adp30684	Human	sec
315	18	60.0	1086	8	ADP31629	Adp31629	Human	sec	388	18	60.0	1506	8	ADP30596	Adp30596	Human	sec
316	18	60.0	1090	8	ADQ10187	Adq10187	Human	pol	389	18	60.0	1510	7	ADG30698	Adg30698	Xanthomon	

390	18	60.0	1518	8	ADP31532	Adp31532	Human	sec	463	18	60.0	1894	6	ABO05229	AbO05229	Novel	hum
391	18	60.0	1530	8	ADP31536	Adp31536	Human	sec	464	18	60.0	1917	6	ADA15719	AdA15719	C. elegans	
392	18	60.0	1539	8	ADP31201	Adp31201	Human	sec	465	18	60.0	1933	8	ADP30889	Adp30889	Human	sec
393	18	60.0	1539	8	ADP31200	Adp31200	Human	sec	466	18	60.0	1933	8	ADP30902	Adp30902	Human	sec
394	18	60.0	1550	8	ADP30567	Adp30567	Human	sec	467	18	60.0	2001	8	ADP31644	Adp31644	Human	sec
395	18	60.0	1584	8	ADP31405	Adp31405	Human	sec	468	18	60.0	2020	8	ADP31056	Adp31056	Human	sec
396	18	60.0	1587	8	ADP30581	Adp30581	Human	sec	469	18	60.0	2020	8	ADP30511	Adp30511	Human	sec
397	18	60.0	1614	8	ADP31529	Adp31529	Human	sec	470	18	60.0	2020	8	ADP30512	Adp30512	Human	sec
398	18	60.0	1617	8	ADP30660	Adp30660	Human	sec	471	18	60.0	2020	8	ADP30509	Adp30509	Human	sec
399	18	60.0	1623	8	ADP30552	Adp30552	Human	sec	472	18	60.0	2027	8	ADP31058	Adp31058	Human	sec
400	18	60.0	1629	8	ADP30945	Adp30945	Human	sec	473	18	60.0	2044	8	ADP30510	Adp30510	Human	sec
401	18	60.0	1647	8	ADP30670	Adp30670	Human	sec	474	18	60.0	2088	8	ADP31178	Adp31178	Human	sec
402	18	60.0	1647	8	ADP30652	Adp30652	Human	sec	475	18	60.0	2091	8	ADP31088	Adp31088	Human	sec
403	18	60.0	1652	6	ADA15715	AdA15715	C. elegans		476	18	60.0	2123	8	ADP30657	Adp30657	Human	sec
404	18	60.0	1665	8	ADP31187	Adp31187	Human	sec	477	18	60.0	2123	8	ADP30564	Adp30564	Human	sec
405	18	60.0	1719	8	ADP31137	Adp31137	Human	sec	478	18	60.0	2127	8	ADP31327	Adp31327	Human	sec
406	18	60.0	1725	8	ADP30654	Adp30654	Human	sec	479	18	60.0	2148	8	ADP30974	Adp30974	Human	sec
407	18	60.0	1743	6	ABU88255	Abu88255	Novel	hum	480	18	60.0	2187	8	ADP30882	Adp30882	Human	sec
408	18	60.0	1743	6	ABU90134	Abu90134	Novel	hum	481	18	60.0	2260	8	ADP30687	Adp30687	Human	sec
409	18	60.0	1743	6	ABU96436	Abu96436	Novel	hum	482	18	60.0	2272	8	ADP30669	Adp30669	Human	sec
410	18	60.0	1743	6	ABU99045	Abu99045	Novel	hum	483	18	60.0	2304	8	ADP31252	Adp31252	Human	sec
411	18	60.0	1743	6	ABU98260	Abu98260	Novel	hum	484	18	60.0	2307	8	ADP31394	Adp31394	Human	sec
412	18	60.0	1743	6	ABU91966	Abu91966	Novel	hum	485	18	60.0	2349	8	ADP30959	Adp30959	Human	sec
413	18	60.0	1743	6	ABU85270	Abu85270	Novel	hum	486	18	60.0	2358	8	ADP31690	Adp31690	Human	sec
414	18	60.0	1743	6	ABO00409	AbO00409	Novel	hum	487	18	60.0	2382	8	ADP31341	Adp31341	Human	sec
415	18	60.0	1743	6	ABU88960	Abu88960	Novel	hum	488	18	60.0	2391	8	ADP31366	Adp31366	Human	sec
416	18	60.0	1743	6	ABO06456	AbO06456	Novel	hum	489	18	60.0	2415	8	ADP31023	Adp31023	Human	sec
417	18	60.0	1743	6	ABU95516	Abu95516	Novel	hum	490	18	60.0	2418	8	ADP31105	Adp31105	Human	sec
418	18	60.0	1743	6	ABU95206	Abu95206	Novel	hum	491	18	60.0	2454	8	ADP30469	Adp30469	Human	sec
419	18	60.0	1743	6	ABU90754	Abu90754	Novel	hum	492	18	60.0	2468	8	ADP30947	Adp30947	Human	sec
420	18	60.0	1743	6	ABU93916	Abu93916	Novel	hum	493	18	60.0	2469	8	ADP31031	Adp31031	Human	sec
421	18	60.0	1743	6	ABU86190	Abu86190	Novel	hum	494	18	60.0	2484	8	ADP66690	Adp66690	Human	mis
422	18	60.0	1743	6	ABU82045	Abu82045	Novel	hum	495	18	60.0	2508	6	ADA15721	AdA15721	C. elegans	
423	18	60.0	1743	6	ABU07906	Abu07906	Novel	hum	496	18	60.0	2514	8	ADP30727	Adp30727	Human	sec
424	18	60.0	1743	6	ABU94226	Abu94226	Novel	hum	497	18	60.0	2535	8	ADP31146	Adp31146	Human	sec
425	18	60.0	1743	6	ABO00099	AbO00099	Novel	hum	498	18	60.0	2544	6	ADA15717	AdA15717	C. elegans	
426	18	60.0	1743	6	ABU87110	Abu87110	Novel	hum	499	18	60.0	2601	6	ADA15723	AdA15723	C. elegans	
427	18	60.0	1743	6	ABU91351	Abu91351	Novel	hum	500	18	60.0	2611	8	ADP31027	Adp31027	Human	sec
428	18	60.0	1743	6	ABU90444	Abu90444	Novel	hum	501	18	60.0	2616	8	ADP31253	Adp31253	Human	sec
429	18	60.0	1743	6	ABU97035	Abu97035	Novel	hum	502	18	60.0	2616	9	ABE849675	AbE849675	N. mening	
430	18	60.0	1743	6	ABO05231	AbO05231	Novel	hum	503	18	60.0	2664	8	ADP31546	Adp31546	Human	sec
431	18	60.0	1755	8	ADP31446	Adp31446	Human	sec	504	18	60.0	2688	8	ADN11593	Adn11593	Human	CD9
432	18	60.0	1766	9	ADP62715	Adp62715	Human	alp	505	18	60.0	2808	8	ADP31258	Adp31258	Human	sec
433	18	60.0	1782	8	ADP31391	Adp31391	Human	sec	506	18	60.0	2832	8	ADP31121	Adp31121	Human	sec
434	18	60.0	1782	8	ADP31270	Adp31270	Human	sec	507	18	60.0	2835	8	ADP30572	Adp30572	Human	sec
435	18	60.0	1789	8	ADP30562	Adp30562	Human	sec	508	18	60.0	2840	8	ADP31120	Adp31120	Human	sec
436	18	60.0	1803	8	ADP30562	Adp30562	Human	sec	509	18	60.0	2980	9	AEA36049	AEa36049	Maize	Str
437	18	60.0	1837	8	ADP31170	Adp31170	Human	sec	510	18	60.0	3046	8	ADQ10419	Adq10419	Human	po1
438	18	60.0	1837	8	ADP30642	Adp30642	Human	sec	511	18	60.0	3070	8	ADP30710	Adp30710	Human	sec
439	18	60.0	1848	8	ADP31372	Adp31372	Human	sec	512	18	60.0	3070	8	ADP30700	Adp30700	Human	sec
440	18	60.0	1894	6	ABU88253	Abu88253	Novel	hum	513	18	60.0	3144	8	ADP31544	Adp31544	Human	sec
441	18	60.0	1894	6	ABU90132	Abu90132	Novel	hum	514	18	60.0	3201	8	ADP31545	Adp31545	Human	sec
442	18	60.0	1894	6	ABU96434	Abu96434	Novel	hum	515	18	60.0	3316	8	ADP31116	Adp31116	Human	sec
443	18	60.0	1894	6	ABU99043	Abu99043	Novel	hum	516	18	60.0	3339	8	ADP31219	Adp31219	Human	sec
444	18	60.0	1894	6	ABU98258	Abu98258	Novel	hum	517	18	60.0	3390	8	ADP31148	Adp31148	Human	sec
445	18	60.0	1894	6	ABU91964	Abu91964	Novel	hum	518	18	60.0	3398	9	ABE87634	AbE87634	Human	ino
446	18	60.0	1894	6	ABU85268	Abu85268	Novel	hum	519	18	60.0	3411	8	ADP30667	Adp30667	Human	sec
447	18	60.0	1894	6	ABO00407	AbO00407	Novel	hum	520	18	60.0	3447	8	ADP31112	Adp31112	Human	sec
448	18	60.0	1894	6	ABU88958	Abu88958	Novel	hum	521	18	60.0	3465	8	ADP31234	Adp31234	Human	sec
449	18	60.0	1894	6	ABO06454	AbO06454	Novel	hum	522	18	60.0	3477	8	ADP30704	Adp30704	Human	sec
450	18	60.0	1894	6	ABU95514	Abu95514	Novel	hum	523	18	60.0	3579	8	ADP31098	Adp31098	Human	sec
451	18	60.0	1894	6	ABU95204	Abu95204	Novel	hum	524	18	60.0	3585	8	ADP31117	Adp31117	Human	sec
452	18	60.0	1894	6	ABU90752	Abu90752	Novel	hum	525	18	60.0	4440	6	ABU88256	Abu88256	Novel	hum
453	18	60.0	1894	6	ABU93914	Abu93914	Novel	hum	526	18	60.0	4440	6	ABU90135	Abu90135	Novel	hum
454	18	60.0	1894	6	ABU86188	Abu86188	Novel	hum	527	18	60.0	4440	6	ABU96437	Abu96437	Novel	hum
455	18	60.0	1894	6	ABU82043	Abu82043	Novel	hum	528	18	60.0	4440	6	ABU99046	Abu99046	Novel	hum
456	18	60.0	1894	6	ABU07904	Abu07904	Novel	hum	529	18	60.0	4440	6	ABU98261	Abu98261	Novel	hum
457	18	60.0	1894	6	ABU94224	Abu94224	Novel	hum	530	18	60.0	4440	6	ABU91967	Abu91967	Novel	hum
458	18	60.0	1894	6	ABO00097	AbO00097	Novel	hum	531	18	60.0	4440	6	ABU85271	Abu85271	Novel	hum
459	18	60.0	1894	6	ABU87108	Abu87108	Novel	hum	532	18	60.0	4440	6	ABO00410	AbO00410	Novel	hum
460	18	60.0	1894	6	ABU91349	Abu91349	Novel	hum	533	18	60.0	4440	6	ABU88961	Abu88961	Novel	hum
461	18	60.0	1894	6	ABU90442	Abu90442	Novel	hum	534	18	60.0	4440	6	ABO06457	AbO06457	Novel	hum
462	18	60.0	1894	6	ABU97033	Abu97033	Novel	hum	535	18	60.0	4440	6	ABU95517	Abu95517	Novel	hum

536	18	60.0	4440	6	ABU95207	Novel	hum	Abu95207	Novel	hum	609	17	56.7	53	6	ABM54353	Abm54353	Proprionib
537	18	60.0	4440	6	ABU90755	Novel	hum	Abu90755	Novel	hum	610	17	56.7	53	6	ABM53320	Abm53320	Proprionib
538	18	60.0	4440	6	ABU93917	Novel	hum	Abu93917	Novel	hum	611	17	56.7	54	6	AEA30185	Aea30185	Pertussis
539	18	60.0	4440	6	ABU86191	Novel	hum	Abu86191	Novel	hum	612	17	56.7	58	4	AAUS4758	Aau54758	Proprionib
540	18	60.0	4440	6	ABU82046	Novel	hum	Abu82046	Novel	hum	613	17	56.7	58	6	ABM51277	Abm51277	Proprionib
541	18	60.0	4440	6	ABU07907	Novel	hum	Abu07907	Novel	hum	614	17	56.7	58	6	AD211761	Ad211761	E. coli D
542	18	60.0	4440	6	ABU94227	Novel	hum	Abu94227	Novel	hum	615	17	56.7	58	9	AEA30433	Aea30433	Pertussis
543	18	60.0	4440	6	ABO00100	Novel	hum	Abu000100	Novel	hum	616	17	56.7	58	9	AEA30308	Aea30308	Pertussis
544	18	60.0	4440	6	ABU87111	Novel	hum	Abu87111	Novel	hum	617	17	56.7	59	4	AAUS1154	Aau51154	Proprionib
545	18	60.0	4440	6	ABU91352	Novel	hum	Abu91352	Novel	hum	618	17	56.7	59	6	ABM47673	Abm47673	Proprionib
546	18	60.0	4440	6	ABU90445	Novel	hum	Abu90445	Novel	hum	619	17	56.7	60	9	AEA30430	Aea30430	Pertussis
547	18	60.0	4440	6	ABU97036	Novel	hum	Abu97036	Novel	hum	620	17	56.7	60	9	AEA30259	Aea30259	Pertussis
548	18	60.0	4440	6	ABO05232	Novel	hum	Abu005232	Novel	hum	621	17	56.7	60	9	AEA30493	Aea30493	Pertussis
549	18	60.0	4683	8	ADP31260	Human sec	Adp31260	Human sec	Adp31260	Human sec	622	17	56.7	60	9	AEA30376	Aea30376	Pertussis
550	18	60.0	4752	8	ADP30585	Human sec	Adp30585	Human sec	Adp30585	Human sec	623	17	56.7	60	9	AEA30494	Aea30494	Pertussis
551	18	60.0	4752	8	ADP30651	Human sec	Adp30651	Human sec	Adp30651	Human sec	624	17	56.7	60	9	AEA30431	Aea30431	Pertussis
552	18	60.0	4838	8	ADP31259	Human sec	Adp31259	Human sec	Adp31259	Human sec	625	17	56.7	60	9	AEA30432	Aea30432	Pertussis
553	18	60.0	5304	8	ADP30706	Human sec	Adp30706	Human sec	Adp30706	Human sec	626	17	56.7	62	7	ADJ38115	Adj38115	Human par
554	18	60.0	5397	8	ADP31068	Human sec	Adp31068	Human sec	Adp31068	Human sec	627	17	56.7	62	7	ABO69433	Abu69433	Pseudomon
555	18	60.0	5514	8	ADP31186	Human sec	Adp31186	Human sec	Adp31186	Human sec	628	17	56.7	64	2	AAV12328	Aay12328	Human 5'
556	18	60.0	5514	8	ADP31591	Human sec	Adp31591	Human sec	Adp31591	Human sec	629	17	56.7	64	2	AAW67942	Aaw67942	Fragment
557	18	60.0	5820	8	ADP31118	Human sec	Adp31118	Human sec	Adp31118	Human sec	630	17	56.7	64	2	AAW65119	Aaw65119	Human imm
558	18	60.0	6465	8	ADP30705	Human sec	Adp30705	Human sec	Adp30705	Human sec	631	17	56.7	64	5	ABP69086	Abp69086	Human pol
559	18	60.0	6729	8	ADP31600	Human sec	Adp31600	Human sec	Adp31600	Human sec	632	17	56.7	66	4	ABG13638	Abg13638	Novel hum
560	18	60.0	7285	6	ABJ38280	pAMG21 -RA	Abj38280	pAMG21 -RA	Abj38280	pAMG21 -RA	633	17	56.7	67	7	ABM86677	Abm86677	Rice abio
561	18	60.0	7339	8	AAO16358	Human tra	Aao16358	Human tra	Aao16358	Human tra	634	17	56.7	68	7	ABP09377	Abp09377	Human ORF
562	18	60.0	8973	8	ADP31119	Human sec	Adp31119	Human sec	Adp31119	Human sec	635	17	56.7	68	7	ABO78891	Abu78891	Pseudomon
563	18	60.0	8976	8	ADP31425	Human sec	Adp31425	Human sec	Adp31425	Human sec	636	17	56.7	70	4	AAUS3345	Aau53345	Proprionib
564	18	60.0	9195	8	ADP31494	Human sec	Adp31494	Human sec	Adp31494	Human sec	637	17	56.7	70	6	ABM49864	Abm49864	Proprionib
565	18	60.0	10944	8	ADP31311	Human sec	Adp31311	Human sec	Adp31311	Human sec	638	17	56.7	71	2	AAW83934	Aaw83934	Human sec
566	18	60.0	11328	8	ADP31310	Human sec	Adp31310	Human sec	Adp31310	Human sec	639	17	56.7	71	3	AAAB34715	Aab34715	Human sec
567	17	56.7	8	5	ABBA47073	Desmocoll	Abba47073	Desmocoll	Abba47073	Desmocoll	640	17	56.7	73	7	ABO82698	Abu82698	Pseudomon
568	17	56.7	8	5	ABBA6610	Desmocoll	Abba6610	Desmocoll	Abba6610	Desmocoll	641	17	56.7	73	4	ADP29540	Adp29540	Human sec
569	17	56.7	8	9	ADZ50990	Amino aci	Adz50990	Amino aci	Adz50990	Amino aci	642	17	56.7	74	4	AAUS9878	Aau59878	Proprionib
570	17	56.7	8	9	ADZ50974	Amino aci	Adz50974	Amino aci	Adz50974	Amino aci	643	17	56.7	74	6	ABM56397	Abm56397	Proprionib
571	17	56.7	10	4	AAAG95975	Human com	Aag95975	Human com	Aag95975	Human com	644	17	56.7	77	9	ADV76639	Adv76639	Human TER
572	17	56.7	10	4	AAAG95917	Human com	Aag95917	Human com	Aag95917	Human com	645	17	56.7	78	8	ADX73420	Adx73420	Plant ful
573	17	56.7	20	2	AAV23344	Peptide 8	Aav23344	Peptide 8	Aav23344	Peptide 8	646	17	56.7	78	9	ABM94485	Abm94485	Human
574	17	56.7	20	3	AAAB19697	O-MechylC	Aab19697	O-MechylC	Aab19697	O-MechylC	647	17	56.7	79	4	AAAI19029	Aai19029	Peptide #
575	17	56.7	20	4	AAAB05831	Liquidamb	Aae05831	Liquidamb	Aae05831	Liquidamb	648	17	56.7	79	4	ABR38235	Abb38235	Peptide #
576	17	56.7	20	9	ADZ14872	Liquidamb	Adz14872	Liquidamb	Adz14872	Liquidamb	649	17	56.7	79	4	AAW31668	Aaw31668	Peptide #
577	17	56.7	22	4	AAW33651	Peptide #	Aam33651	Peptide #	Aam33651	Peptide #	650	17	56.7	79	4	ABR32413	Abb32413	Protein #
578	17	56.7	22	4	AAW73453	Human bon	Aam73453	Human bon	Aam73453	Human bon	651	17	56.7	79	4	ABG53088	Abg53088	Human liv
579	17	56.7	22	4	AAW60779	Human bra	Aam60779	Human bra	Aam60779	Human bra	652	17	56.7	79	5	ABG41186	Abg41186	Human pep
580	17	56.7	22	4	ABG55176	Human liv	Abg55176	Human liv	Abg55176	Human liv	653	17	56.7	80	3	AAAI1389	Aai1389	Proprionib
581	17	56.7	22	5	ABG43312	Human pep	Abg43312	Human pep	Abg43312	Human pep	654	17	56.7	80	4	AAU64692	Aau64692	Proprionib
582	17	56.7	24	8	ADP73753	Influenza	Adp73753	Influenza	Adp73753	Influenza	655	17	56.7	83	6	ABM61211	Abm61211	Proprionib
583	17	56.7	28	4	AAAG98776	Human cel	Aag98776	Human cel	Aag98776	Human cel	656	17	56.7	84	3	AAAG60503	Aag60503	Arabidops
584	17	56.7	28	9	ADZ47463	Human HAS	Adz47463	Human HAS	Adz47463	Human HAS	657	17	56.7	84	3	AAAG55303	Aag55303	Arabidops
585	17	56.7	34	5	AAU87022	Adenoviru	Aau87022	Adenoviru	Aau87022	Adenoviru	658	17	56.7	84	4	AAU61337	Aau61337	Proprionib
586	17	56.7	35	6	ABU61319	Human A d	Abu61319	Human A d	Abu61319	Human A d	659	17	56.7	84	6	ABM57856	Abm57856	Proprionib
587	17	56.7	35	6	ABU61408	Low densi	Abu61408	Low densi	Abu61408	Low densi	660	17	56.7	85	9	ABR41187	Abb41187	Proprionib
588	17	56.7	35	8	ADP21541	Human LDL	Adp21541	Human LDL	Adp21541	Human LDL	661	17	56.7	86	4	AAU51874	Aau51874	Proprionib
589	17	56.7	37	4	AAO09823	Human pol	Aao09823	Human pol	Aao09823	Human pol	662	17	56.7	86	6	ABM48393	Abm48393	Proprionib
590	17	56.7	39	9	ABR43285	Human LRP	Aeb43285	Human LRP	Aeb43285	Human LRP	663	17	56.7	87	4	AAU63689	Aau63689	Proprionib
591	17	56.7	50	6	ABM56396	Proprionib	Abm56396	Proprionib	Abm56396	Proprionib	664	17	56.7	87	6	ABM60208	Abm60208	Proprionib
592	17	56.7	50	6	ABM56396	Proprionib	Abm56396	Proprionib	Abm56396	Proprionib	665	17	56.7	87	7	ADP17451	Adp17451	Mouse IL-
593	17	56.7	51	4	ABR39045	Peptide #	Aam39045	Peptide #	Aam39045	Peptide #	666	17	56.7	88	4	AAU21754	Aau21754	Novel hum
594	17	56.7	51	4	AAW32534	Peptide #	Aam32534	Peptide #	Aam32534	Peptide #	667	17	56.7	88	4	AAU45755	Aau45755	Proprionib
595	17	56.7	51	4	AAW2275	Human bon	Aam2275	Human bon	Aam2275	Human bon	668	17	56.7	88	4	AAU20119	Aau20119	Proprionib
596	17	56.7	51	4	AAW59697	Human bra	Aam59697	Human bra	Aam59697	Human bra	669	17	56.7	88	5	ABG91368	Abg91368	Novel hum
597	17	56.7	51	4	ABG53961	Human bra	Abg53961	Human bra	Abg53961	Human bra	670	17	56.7	88	6	ABM42274	Abm42274	Proprionib
598	17	56.7	51	5	ABG42090	Human pep	Abg42090	Human pep	Abg42090	Human pep	671	17	56.7	88	7	ADC46395	Adc46395	Human neo
599	17	56.7	53	4	AAW19878	Peptide #	Aam19878	Peptide #	Aam19878	Peptide #	672	17	56.7	88	7	ABO69228	Abu69228	Pseudomon
600	17	56.7	53	4	ABR39811	Peptide #	Abb39811	Peptide #	Abb39811	Peptide #	673	17	56.7	90	8	ADP31655	Adp31655	Human sec
601	17	56.7	53	4	AAW33401	Protein #	Abb33401	Protein #	Abb33401	Protein #	674	17	56.7	91	2	AAV12296	Aay12296	Human 5'
602	17	56.7	53	4	ABR24424	Protein #	Abb24424	Protein #	Abb24424	Protein #	675	17	56.7	92	5	ABO09944	Abu09944	Guinea pi
603	17	56.7	53	4	AAW73191	Human bon	Aam73191	Human bon	Aam73191	Human bon	676	17	56.7	92	5	ABO09944	Abu09944	Guinea pi
604	17	56.7	53	4	AAUS6801	Proprionib	Aau6801	Proprionib	Aau6801	Proprionib	677	17	56.7	92	7	AAE39395	Aae39395	Human int
605	17	56.7	53	4	AAUS7834	Proprionib	Aau7834	Proprionib	Aau7834	Proprionib	678	17	56.7	92	7	AAE39395	Aae39395	Human int
606	17	56.7	53	4	AAW60535	Human bra	Aam60535	Human bra	Aam60535	Human bra	679	17	56.7	92	9	AEA35800	Aea35800	Partial i
607	17	56.7	53	4	ABG54907	Human liv	Abg54907	Human liv	Abg54907	Human liv	680	17	56.7	92	9	AEA35802	Aea35802	Partial i
608	17	56.7	53	5	ABG43038	Human pep	Abg43038	Human pep	Abg43038	Human pep	681	17	56.7	93	4	AAU56711	Aau56711	Proprionib

682	17	56.7	93	6	ABMS3230	Abm53230	Propionib	755	17	56.7	134	6	ABMS5230	Abm55230	Propionib
683	17	56.7	96	4	AAUS2513	Aau52513	Propionib	756	17	56.7	134	7	ABO76893	AbO76893	Pseudomon
684	17	56.7	96	6	ABM49032	Abm49032	Propionib	757	17	56.7	135	3	AAV75297	Aay75297	Neisseria
685	17	56.7	97	4	AAU66144	Aau66144	Propionib	758	17	56.7	135	3	AAV75296	Aay75296	Neisseria
686	17	56.7	97	5	ABP09938	Abp09938	Human ORF	759	17	56.7	135	5	ABP08959	Abp08959	Human ORF
687	17	56.7	97	6	ABM62663	Abm62663	Human ORF	760	17	56.7	135	5	ABO74109	AbO74109	Pseudomon
688	17	56.7	97	6	ABO79445	AbO79445	Pseudomon	761	17	56.7	135	8	ADP31538	Adp31538	Human sec
689	17	56.7	97	9	ABE37882	Aeb37882	L. pneumo	762	17	56.7	136	8	ADY07492	Ady07492	Plant full
690	17	56.7	99	4	AAU49282	Aau49282	Propionib	763	17	56.7	136	3	AAV75295	Aay75295	Neisseria
691	17	56.7	99	5	ABP43926	Abp43926	NICE-1 pr	764	17	56.7	136	7	ADMS7545	Adm57545	Human fls
692	17	56.7	99	6	ABM45781	Abm45781	Propionib	765	17	56.7	136	7	ABO73988	AbO73988	Pseudomon
693	17	56.7	99	8	ADN04970	Adn04970	Antipsoxi	766	17	56.7	136	7	ABM66687	Abm66687	Rice abio
694	17	56.7	102	7	ABO79727	AbO79727	Pseudomon	767	17	56.7	137	3	AAV65413	Aay65413	Human 5'
695	17	56.7	105	7	ADBE2627	Adbe2627	Rat Prote	768	17	56.7	137	8	ADU72977	Adu72977	Signal pe
696	17	56.7	106	4	AAAG4065	Aag4065	Human ant	769	17	56.7	137	8	ADZ73968	Adz73968	Human com
697	17	56.7	106	4	AAAG4068	Aag4068	Rat anter	770	17	56.7	140	7	ADFL13949	Adfl13949	Human end
698	17	56.7	106	4	ABBI5333	Abbi5333	Human nar	771	17	56.7	140	7	ABO82130	AbO82130	Pseudomon
699	17	56.7	106	5	AAU96154	Aau96154	Human nar	772	17	56.7	141	7	ABO77083	AbO77083	Pseudomon
700	17	56.7	106	6	ADA57463	Ada57463	Human sec	773	17	56.7	143	5	AAU96140	Aau96140	Human BRP
701	17	56.7	106	6	ADA41343	Ada41343	Human sec	774	17	56.7	144	7	ABO79215	AbO79215	Pseudomon
702	17	56.7	106	6	ABR48065	AbR48065	Human sec	775	17	56.7	144	8	ADJ379724	Adj379724	TADG-14 p
703	17	56.7	106	6	ABP72423	Abp72423	Human gly	776	17	56.7	144	8	ADP31474	Adp31474	Human pro
704	17	56.7	106	8	ADF72484	Adf72484	Mature hu	777	17	56.7	144	8	ADP31474	Adp31474	Human sec
705	17	56.7	106	9	ADM86161	Adm86161	Human Zlu	778	17	56.7	148	7	ADP13954	Adp13954	Human end
706	17	56.7	107	7	ADP13950	Adp13950	Human end	779	17	56.7	148	7	ABO74263	AbO74263	Pseudomon
707	17	56.7	107	8	ADP30723	Adp30723	Human sec	780	17	56.7	149	4	AAU68926	Aau68926	Human pro
708	17	56.7	111	4	ABG03591	Abg03591	Novel hum	781	17	56.7	149	7	ADP31059	Adp31059	Human dia
709	17	56.7	112	7	ABO77337	AbO77337	Pseudomon	782	17	56.7	149	7	ABE13246	AbE13246	Protease
710	17	56.7	113	4	AAUS0472	Aau50472	Novel hum	783	17	56.7	149	8	ADH78455	Adh78455	Human pro
711	17	56.7	113	6	ABM46991	Abm46991	Propionib	784	17	56.7	150	8	ADK65524	Adk65524	Plant full
712	17	56.7	113	8	ADP31617	Adp31617	Human sec	785	17	56.7	154	7	ABO74102	AbO74102	Pseudomon
713	17	56.7	113	8	ADP31613	Adp31613	Human sec	786	17	56.7	154	8	ABO60522	AbO60522	Human gen
714	17	56.7	113	3	AAAB25477	Aab25477	Pinus rad	787	17	56.7	154	7	ABO80146	AbO80146	Pseudomon
715	17	56.7	117	7	ADBE5158	Adbe5158	Human pro	788	17	56.7	156	8	ADP30759	Adp30759	Human sec
716	17	56.7	117	7	ABO68458	AbO68458	Pseudomon	789	17	56.7	157	8	ADK78826	Adk78826	Plant full
717	17	56.7	118	6	ABU99121	Abu99121	Novel hum	790	17	56.7	162	3	AAV95932	Aay95932	Porcine a
718	17	56.7	118	8	ADM93807	Adm93807	Human NOV	791	17	56.7	162	4	AAU42696	Aau42696	Propionib
719	17	56.7	120	8	ADB64538	Adb64538	Human pro	792	17	56.7	162	6	ABM39215	Abm39215	Propionib
720	17	56.7	120	8	ADP30556	Adp30556	Human sec	793	17	56.7	163	7	ABO80365	AbO80365	Pseudomon
721	17	56.7	122	4	AAU42386	Aau42386	Propionib	794	17	56.7	163	7	ABO68705	AbO68705	Pseudomon
722	17	56.7	122	6	ABM38905	Abm38905	Propionib	795	17	56.7	164	7	ABO78378	AbO78378	Pseudomon
723	17	56.7	123	7	ABO66168	AbO66168	Klebsiell	796	17	56.7	165	4	AAUS8075	Aau58075	Propionib
724	17	56.7	124	4	AAUS1146	Aaus1146	Propionib	797	17	56.7	165	6	ABMS4594	Abms4594	Propionib
725	17	56.7	124	6	ABM47665	Abm47665	Propionib	798	17	56.7	165	8	ADP31286	Adp31286	Human sec
726	17	56.7	127	8	ADP31622	Adp31622	Human sec	799	17	56.7	165	8	ADP31515	Adp31515	Human sec
727	17	56.7	129	4	AAAG4067	Aag4067	Rat anter	800	17	56.7	165	8	ADP31174	Adp31174	Human sec
728	17	56.7	129	5	AAU96157	Aau96157	Human BRP	801	17	56.7	165	8	ADP31516	Adp31516	Human sec
729	17	56.7	129	5	AAU96129	Aau96129	Human bet	802	17	56.7	165	8	ADP31041	Adp31041	Human sec
730	17	56.7	129	6	ABU97054	Abu97054	Recombina	803	17	56.7	168	8	ADP31621	Adp31621	Human sec
731	17	56.7	130	4	AAAG3211	Aag3211	Mutino ac1	804	17	56.7	168	8	ADP30797	Adp30797	Human sec
732	17	56.7	130	4	AAAG4064	Aag4064	Human ant	805	17	56.7	168	8	ADP30769	Adp30769	Human sec
733	17	56.7	130	4	AAU41366	Aau41366	Propionib	806	17	56.7	168	8	ADP30771	Adp30771	Human sec
734	17	56.7	130	4	AAE09440	Aae09440	Human sbg	807	17	56.7	171	8	ADP30599	Adp30599	Human sec
735	17	56.7	130	5	AAU97612	Aau97612	Human OGH	808	17	56.7	175	3	AAAG54788	Aag54788	Arabidops
736	17	56.7	130	5	AAU96153	Aau96153	Human BRP	809	17	56.7	175	7	ABO80411	AbO80411	Pseudomon
737	17	56.7	130	5	AAU10368	Aau10368	Mouse bet	810	17	56.7	175	7	ABO69379	AbO69379	Pseudomon
738	17	56.7	130	5	AAU10366	Aau10366	Human bet	811	17	56.7	177	7	ABO70036	AbO70036	Pseudomon
739	17	56.7	130	6	ABG74215	Abg74215	Mouse gly	812	17	56.7	178	7	ABO74750	AbO74750	Pseudomon
740	17	56.7	130	6	ABG74206	Abg74206	Human gly	813	17	56.7	178	7	ABO73284	AbO73284	Pseudomon
741	17	56.7	130	6	ABM37885	Abm37885	Propionib	814	17	56.7	179	7	ABO75778	AbO75778	Pseudomon
742	17	56.7	130	6	ABP72432	Abp72432	Human gly	815	17	56.7	180	8	ADP30820	Adp30820	Human sec
743	17	56.7	130	7	ADJ38121	Adj38121	Human OGH	816	17	56.7	180	8	ADP30825	Adp30825	Human sec
744	17	56.7	130	7	ABO81191	AbO81191	Pseudomon	817	17	56.7	180	8	ADP30828	Adp30828	Human sec
745	17	56.7	130	8	ADP72483	Adp72483	Human GPH	818	17	56.7	180	8	ADP30821	Adp30821	Human sec
746	17	56.7	130	8	ADM6155	Adm6155	Human Zlu	819	17	56.7	180	8	ADP30826	Adp30826	Human sec
747	17	56.7	130	9	ADM6175	Adm6175	Marine z1	820	17	56.7	180	8	ADP30827	Adp30827	Human sec
748	17	56.7	130	9	AE812349	Aeb12349	Human cyp	821	17	56.7	182	7	ABO78099	AbO78099	Pseudomon
749	17	56.7	131	5	AAU96155	Aau96155	Human BRP	822	17	56.7	182	7	ABO71666	AbO71666	Pseudomon
750	17	56.7	133	4	AAU49190	Aau49190	Propionib	823	17	56.7	182	8	ADN99530	Adn99530	Novel hum
751	17	56.7	133	4	ABG04654	Abg04654	Novel hum	824	17	56.7	184	7	ABO77341	AbO77341	Pseudomon
752	17	56.7	133	6	ABM45709	Abm45709	Propionib	825	17	56.7	185	8	ADP31109	Adp31109	Human sec
753	17	56.7	133	7	ABO82412	AbO82412	Pseudomon	826	17	56.7	186	4	ABG08184	AbG08184	Novel hum
754	17	56.7	134	4	AAUS8711	Aaus8711	Propionib	827	17	56.7	187	7	ABO80088	AbO80088	Pseudomon

828	17	56.7	190	6	ABM65810	Abm65810	Propionib	901	17	56.7	249	3	AA660495	AA660495	Arabidops
829	17	56.7	191	7	ABO76619	ABO76619	Pseudomon	902	17	56.7	249	7	ABO83546	ABO83546	Abop3546
830	17	56.7	192	8	ADP31335	Human sec	903	17	56.7	250	6	ABE65693	ABE65693	E. tenell	
831	17	56.7	192	8	ADP30575	Human sec	904	17	56.7	250	7	ABO82689	ABO82689	Abob2689	
832	17	56.7	198	4	AAU47307	Propionib	905	17	56.7	251	7	ADJ92161	ADJ92161	Human hai	
833	17	56.7	198	6	ABM43826	Abm43826	Propionib	906	17	56.7	251	7	ABO73872	ABO73872	Abob3872
834	17	56.7	198	8	ADP30811	Human sec	907	17	56.7	251	9	ABM91192	ABM91192	Abm91192	
835	17	56.7	198	8	ADP30811	Human sec	908	17	56.7	254	8	ADP30739	ADP30739	Human sec	
836	17	56.7	198	8	ADP30477	Human sec	909	17	56.7	254	8	ADP31396	ADP31396	Adp31396	
837	17	56.7	198	8	ADP30481	Human sec	910	17	56.7	255	7	ADJ92145	ADJ92145	Human hai	
838	17	56.7	200	7	ABO71047	Human sec	911	17	56.7	255	8	ADP30658	ADP30658	Human sec	
839	17	56.7	201	7	ABO60877	AbO60877	Klebsiell	912	17	56.7	256	4	ABG08337	ABG08337	Novel hum
840	17	56.7	203	7	ADB80007	AbB80007	Mycobacte	913	17	56.7	256	6	ABO97138	ABO97138	Abu97138
841	17	56.7	204	8	ADP31404	Human sec	914	17	56.7	256	7	ABO75228	ABO75228	AbO75228	
842	17	56.7	206	7	ABO74022	AbO74022	Pseudomon	915	17	56.7	256	8	ADN99494	ADN99494	Adn99494
843	17	56.7	208	4	AAU29506	AAU29506	Human G p	916	17	56.7	257	5	ABP64573	ABP64573	ABP64573
844	17	56.7	208	5	ABG60794	Novel G p	917	17	56.7	258	8	ADP31084	ADP31084	ADP31084	
845	17	56.7	208	7	ABO80531	AbO80531	Pseudomon	918	17	56.7	259	7	ABM8128	ABM8128	Abm8128
846	17	56.7	209	7	ABO77021	AbO77021	Pseudomon	919	17	56.7	261	6	AAO27174	AAO27174	AAO27174
847	17	56.7	210	8	ADP31563	Human sec	920	17	56.7	263	3	AA660494	AA660494	AA660494	
848	17	56.7	212	7	ABO67532	AbO67532	Klebsiell	921	17	56.7	263	7	ABO77554	ABO77554	AbO77554
849	17	56.7	212	7	ADP31562	ADP31562	Human sec	922	17	56.7	263	8	ABM81215	ABM81215	Abm81215
850	17	56.7	215	7	ADP08235	Novel pro	923	17	56.7	264	8	ADP30788	ADP30788	ADP30788	
851	17	56.7	215	7	ABO72061	AbO72061	Pseudomon	924	17	56.7	264	8	ADP31456	ADP31456	ADP31456
852	17	56.7	215	9	ABM93776	AbM93776	M. xanthu	925	17	56.7	264	8	ADP31412	ADP31412	ADP31412
853	17	56.7	217	4	AAJ97773	AAJ97773	I. scapul	926	17	56.7	264	8	ADP31427	ADP31427	ADP31427
854	17	56.7	218	7	ABO81158	AbO81158	Pseudomon	927	17	56.7	266	8	ADP31427	ADP31427	ADP31427
855	17	56.7	218	8	ADY23534	ADY23534	Plant ful	928	17	56.7	267	4	ABG27981	ABG27981	Novel hum
856	17	56.7	219	8	ADP31171	ADP31171	Human sec	929	17	56.7	267	8	ADP31483	ADP31483	ADP31483
857	17	56.7	219	8	ADY06385	ADY06385	Plant ful	930	17	56.7	270	8	ADP31217	ADP31217	ADP31217
858	17	56.7	221	7	ADJ92147	ADJ92147	Human hai	931	17	56.7	270	8	ADP31321	ADP31321	ADP31321
859	17	56.7	221	8	ADM87372	ADM87372	Human pro	932	17	56.7	270	8	ADP31564	ADP31564	ADP31564
860	17	56.7	222	5	ABM78572	AbM78572	Rat ribon	933	17	56.7	270	9	ABE12365	ABE12365	ABE12365
861	17	56.7	222	5	ABM78573	AbM78573	Mouse rib	934	17	56.7	271	6	ABU34590	ABU34590	ABU34590
862	17	56.7	223	3	AAE03862	AAE03862	Human neu	935	17	56.7	271	6	ABU34590	ABU34590	ABU34590
863	17	56.7	224	2	AAW22018	AAW22018	Erg myoge	936	17	56.7	271	6	ADJ92151	ADJ92151	ADJ92151
864	17	56.7	224	9	ADW95939	ADW95939	Human myo	937	17	56.7	272	6	ABU35898	ABU35898	ABU35898
865	17	56.7	224	9	ADX15805	ADX15805	Human myo	938	17	56.7	272	8	ADP12954	ADP12954	ADP12954
866	17	56.7	225	5	ABM78570	AbM78570	Human rib	939	17	56.7	273	8	ADP31236	ADP31236	ADP31236
867	17	56.7	225	7	ABM88876	AbM88876	Rice abio	940	17	56.7	273	9	ABM94710	ABM94710	ABM94710
868	17	56.7	225	8	ADP31501	ADP31501	Human sec	941	17	56.7	274	4	AAU19418	AAU19418	AAU19418
869	17	56.7	225	8	ADP30777	ADP30777	Human sec	942	17	56.7	274	7	ABO72003	ABO72003	ABO72003
870	17	56.7	227	7	ABO75458	ABO75458	Pseudomon	943	17	56.7	274	8	ADP31679	ADP31679	ADP31679
871	17	56.7	228	8	ADP30921	ADP30921	Human sec	944	17	56.7	275	6	ABU34080	ABU34080	ABU34080
872	17	56.7	231	5	ABP64726	ABP64726	Human pro	945	17	56.7	276	8	ABO75093	ABO75093	ABO75093
873	17	56.7	233	7	ABO75933	ABO75933	Pseudomon	946	17	56.7	276	8	ADP31319	ADP31319	ADP31319
874	17	56.7	235	7	ABO69554	ABO69554	Pseudomon	947	17	56.7	276	8	ADP30584	ADP30584	ADP30584
875	17	56.7	237	8	ADP31322	ADP31322	Human sec	948	17	56.7	278	8	ADP31314	ADP31314	ADP31314
876	17	56.7	238	9	ABE12367	ABE12367	Fusion pr	949	17	56.7	279	8	ADP30846	ADP30846	ADP30846
877	17	56.7	239	3	AAAB21310	AAAB21310	Human zym	950	17	56.7	279	8	ADP31523	ADP31523	ADP31523
878	17	56.7	239	8	ADY13038	ADY13038	Plant ful	951	17	56.7	279	8	ADP31503	ADP31503	ADP31503
879	17	56.7	240	2	AAJ22205	AAJ22205	Biorthytm	952	17	56.7	279	8	ADP31671	ADP31671	ADP31671
880	17	56.7	240	2	ADP30737	ADP30737	Human sec	953	17	56.7	279	8	ADP30605	ADP30605	ADP30605
881	17	56.7	244	2	AAK44532	AAK44532	Yyme App-	954	17	56.7	279	8	ADP30844	ADP30844	ADP30844
882	17	56.7	244	2	AAW22985	AAW22985	Human ser	955	17	56.7	282	7	ADJ92143	ADJ92143	ADJ92143
883	17	56.7	244	2	AAW51006	AAW51006	Protease	956	17	56.7	285	7	ABO73521	ABO73521	ABO73521
884	17	56.7	244	3	AAAB21323	AAAB21323	Human zym	957	17	56.7	288	8	ADP31463	ADP31463	ADP31463
885	17	56.7	244	5	ABG96357	ABG96357	Human ova	958	17	56.7	291	7	ABO68749	ABO68749	ABO68749
886	17	56.7	244	6	AAE37572	AAE37572	Human 204	959	17	56.7	293	6	ADN55092	ADN55092	ADN55092
887	17	56.7	244	7	ADB80567	ADB80567	Ovarian c	960	17	56.7	293	8	ADH09610	ADH09610	ADH09610
888	17	56.7	244	7	ADN39212	ADN39212	Cancer/an	961	17	56.7	294	8	ADP31076	ADP31076	ADP31076
889	17	56.7	244	8	ADJ39734	ADJ39734	Human pro	962	17	56.7	294	8	ADP31473	ADP31473	ADP31473
890	17	56.7	244	8	ADJ37158	ADJ37158	Human pro	963	17	56.7	297	8	ADP31192	ADP31192	ADP31192
891	17	56.7	244	8	ADN04074	ADN04074	Antipsoi	964	17	56.7	298	2	AAW43397	AAW43397	AAW43397
892	17	56.7	244	8	ADN29289	ADN29289	Human kal	965	17	56.7	302	3	AAO07023	AAO07023	AAO07023
893	17	56.7	244	8	ADQ89076	ADQ89076	Human uro	966	17	56.7	304	8	ADP31659	ADP31659	ADP31659
894	17	56.7	244	8	ADR72624	ADR72624	Human ren	967	17	56.7	306	7	ADFO4021	ADFO4021	ADFO4021
895	17	56.7	244	8	ADR72876	ADR72876	Human ova	968	17	56.7	307	7	ADCO1287	ADCO1287	ADCO1287
896	17	56.7	244	9	ADY67594	ADY67594	Human kal	969	17	56.7	309	8	ADJ42146	ADJ42146	ADJ42146
897	17	56.7	244	9	ADZ51357	ADZ51357	Amino aci	970	17	56.7	310	4	ABB58993	ABB58993	ABB58993
898	17	56.7	245	7	ABO78304	ABO78304	Pseudomon	971	17	56.7	311	4	ABE67853	ABE67853	ABE67853
899	17	56.7	246	8	ADP30619	ADP30619	Human sec	972	17	56.7	311	4	ABG08405	ABG08405	Novel hum
900	17	56.7	246	9	ADY18495	ADY18495	PRO polyP	973	17	56.7	312	1	AAE50079	AAE50079	T-cell an

```
974 17 56.7 312 1 AAP60471 Aap60471 Portion o
975 17 56.7 312 2 AAR53145 Aar53145 T-cell an
976 17 56.7 312 3 AAB25555 Aab25555 Pinus rad
977 17 56.7 312 4 ABG08404 Abg08404 Novel hum
978 17 56.7 312 8 ADP31505 Adp31505 Human sec
979 17 56.7 312 8 ADP30476 Adp30476 Human sec
980 17 56.7 313 7 ABM85966 Abm85966 Rice abio
981 17 56.7 315 5 AAE13348 Aae13348 Human 1ST
982 17 56.7 315 6 ABR58684 Abr58684 Human can
983 17 56.7 315 8 ADP31685 Adp31685 Human tas
984 17 56.7 316 3 AAY44862 Aay44862 Human sec
985 17 56.7 316 9 ADY18921 Ady18921 PRO polyP
986 17 56.7 317 8 ADP31615 Adp31615 Human sec
987 17 56.7 318 7 ABO78573 Abo78573 Pseudomon
988 17 56.7 318 8 ADP31066 Adp31066 Human sec
989 17 56.7 320 4 AAG91631 Aag91631 C glutamI
990 17 56.7 325 3 AAB51703 Aab51703 Gene 34 h
991 17 56.7 326 3 AAG24798 Aag24798 Arabidops
992 17 56.7 336 8 ADP31320 Adp31320 Human sec
993 17 56.7 341 7 ABO77462 Abo77462 Pseudomon
994 17 56.7 343 7 ADP13953 Adp13953 Human end
995 17 56.7 345 3 AAG07022 Aag07022 Arabidops
996 17 56.7 345 3 AAG37538 Aag37538 Arabidops
997 17 56.7 345 8 ADP31204 Adp31204 Human sec
998 17 56.7 346 7 ABO81210 Abo81210 Pseudomon
999 17 56.7 346 7 ABM89803 Abm89803 Rice abio
1000 17 56.7 346 8 ADP31355 Adp31355 Human sec
```

## ALIGNMENTS

```
RESULT 1
ADQ13022
ID ADQ13022 standard; peptide; 9 AA.
XX
AC ADQ13022;
XX
DT 07-OCT-2004 (first entry)
XX
DE Hepatitis B virus MHC class I restricted T-cell stimulating peptide 821.
XX
KM HBV, MHC class I restricted T-cell stimulating peptide; surface peptide;
KW core peptide; polymerase peptide; immunogenic composition; vaccine;
XX cytotoxic T-lymphocyte response; CTL response.
XX OS Hepatitis B virus.
XX PN WO2004058807-A2.
XX
XX PD 15-JUL-2004.
XX
XX PF 09-DEC-2003; 2003WO-EP013948.
XX
XX PR 24-DEC-2002; 2002EP-00447276.
XX
XX PA (ALGO-) ALGONOMICS NV.
XX
XX PI Lasters I, Desmet J, Stegmann T, Castelein B;
XX
XX DR WPI, 2004-525861/50.
XX
XX PT New peptide comprising a major histocompatibility complex class I
XX restricted T-cell stimulating epitope of the hepatitis B virus (HBV)
XX surface, core and/or polymerase polypeptide, for preparation of an HBV
XX immunogenic composition.
XX
XX PS Claim 5; SEQ ID NO 821; 108bp; English.
XX
XX CC The invention comprises major histocompatibility complex (MHC) class I
XX restricted T-cell stimulating epitopes of the hepatitis B virus (HBV)
XX surface, core and polymerase proteins. The peptides of the invention are
XX useful for generating an HBV immunogenic composition (e.g. vaccine) that
```

CC induces a cytotoxic T-lymphocyte (CTL) response. The present amino acid  
CC sequence represents an MHC class I restricted T-cell stimulating HBV  
CC peptide of the invention.

SQ Sequence 9 AA;

Query Match 60.0%; Score 18; DB 8; Length 9;  
Best Local Similarity 25.0%; Pred. No. 1.4e+06;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 1 CASSSSSC 8

```
RESULT 2
ADQ13021
ID ADQ13021 standard; peptide; 9 AA.
```

```
XX
AC ADQ13021;
XX
DT 07-OCT-2004 (first entry)
```

DE Hepatitis B virus MHC class I restricted T-cell stimulating peptide 820.

KM HBV, MHC class I restricted T-cell stimulating peptide; surface peptide;  
KW core peptide; polymerase peptide; immunogenic composition; vaccine;  
XX cytotoxic T-lymphocyte response; CTL response.

OS Hepatitis B virus.

PN WO2004058807-A2.

XX PD 15-JUL-2004.

XX PF 09-DEC-2003; 2003WO-EP013948.

XX PR 24-DEC-2002; 2002EP-00447276.

XX PA (ALGO-) ALGONOMICS NV.

XX PI Lasters I, Desmet J, Stegmann T, Castelein B;

XX DR WPI, 2004-525861/50.

XX PT New peptide comprising a major histocompatibility complex class I  
XX restricted T-cell stimulating epitope of the hepatitis B virus (HBV)  
XX surface, core and/or polymerase polypeptide, for preparation of an HBV  
XX immunogenic composition.

XX PS Example 7; SEQ ID NO 820; 108bp; English.

XX CC The invention comprises major histocompatibility complex (MHC) class I  
XX restricted T-cell stimulating epitopes of the hepatitis B virus (HBV)  
XX surface, core and polymerase proteins. The peptides of the invention are  
XX useful for generating an HBV immunogenic composition (e.g. vaccine) that  
XX induces a cytotoxic T-lymphocyte (CTL) response. The present amino acid  
XX sequence represents an MHC class I restricted T-cell stimulating HBV  
XX peptide of the invention.

XX SQ Sequence 9 AA;

Query Match 60.0%; Score 18; DB 8; Length 9;  
Best Local Similarity 25.0%; Pred. No. 1.4e+06;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 2 CASSSSSC 9

```
RESULT 3
AAW65943
```



```

ID AAW65943 standard; peptide; 11 AA.
XX
AC AAW65943;
XX
DT 12-NOV-1998 (first entry)
XX
DE Molecule VI.
XX
KM Neurotrophin; cyclic; bicyclic; disulphide; nerve growth factor; NGF;
XX BDNF; NT-3; conformation; promoter.
XX
OS Synthetic.
OS Mammalia.
XX
PN CA2205045-A.
XX
PD 12-MAY-1998.
XX
PF 09-MAY-1997; 97CA-02205045.
XX
PR 12-NOV-1996; 96CA-02190296.
XX
PA (TOOH ) UNIV QUEBENS KINGSTON.
XX
PI Riopelle RJ, Dory MI, Shamovsky IL, Weaver DF, Ross GM;
XX
DR WPI; 1998-457650/40.
XX
PT Inhibition of neurotrophin activity - using factor that interferes with
XX sub-unit interaction.
XX
PS Disclosure; Page 13; 103pp; English.
XX
CC The invention relates to a method for reducing the biological activity of
CC a multimeric protein having at least 2 promoters. It comprises (a)
CC providing a factor that interacts with at least one portion of at least
CC one of the promoters which associates with a portion of the other
CC promoter in the absence of the factor; and (b) mixing the factor with the
CC multimeric protein so that the factor interacts with the portion(s) and
CC disrupts association of at least a portion of the promoters. The method
CC and compounds are useful for inhibiting neurotrophin-mediated activities
CC selected from neurotrophin receptor binding, neuron survival, neurite
CC outgrowth and epileptic effects. The present sequence is shown in the
CC specification
XX
SQ Sequence 11 AA;
XX
Query Match 60.0%; Score 18; DB 2; Length 11;
Best Local Similarity 25.0%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 1 CAAAAAAC 8

```

RESULT 4  
ID AAW65944 standard; peptide; 11 AA.  
XX  
AC AAW65944;  
XX  
DT 12-NOV-1998 (first entry)  
XX  
DE Molecule XV.  
XX  
KM Neurotrophin; cyclic; bicyclic; disulphide; nerve growth factor; NGF;  
XX BDNF; NT-3; conformation; promoter.  
XX  
OS Synthetic.  
OS Mammalia.  
XX  
PN CA2205045-A.

```

XX
XX 12-MAY-1998.
XX
PD 09-MAY-1997; 97CA-02205045.
XX
PF 12-NOV-1996; 96CA-02190296.
XX
PR (TOOH ) UNIV QUEBENS KINGSTON.
XX
PA Riopelle RJ, Dory MI, Shamovsky IL, Weaver DF, Ross GM;
XX
PI WPI; 1998-457650/40.
XX
DR Inhibition of neurotrophin activity - using factor that interferes with
XX sub-unit interaction.
XX
PT Disclosure; Page 13; 103pp; English.
XX
PS The invention relates to a method for reducing the biological activity of
XX a multimeric protein having at least 2 promoters. It comprises (a)
XX providing a factor that interacts with at least one portion of at least
XX one of the promoters which associates with a portion of the other
XX promoter in the absence of the factor; and (b) mixing the factor with the
XX multimeric protein so that the factor interacts with the portion(s) and
XX disrupts association of at least a portion of the promoters. The method
XX and compounds are useful for inhibiting neurotrophin-mediated activities
XX selected from neurotrophin receptor binding, neuron survival, neurite
XX outgrowth and epileptic effects. The present sequence is shown in the
XX specification
XX
SQ Sequence 11 AA;
XX
Query Match 60.0%; Score 18; DB 2; Length 11;
Best Local Similarity 25.0%; Pred. No. 2.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CXXXXXXC 11
DB 4 CAAAAAAC 11

```

RESULT 5  
ID ADV23688 standard; peptide; 15 AA.  
XX  
AC ADV23688;  
XX  
DT 10-MAR-2005 (first entry)  
XX  
DE HBV immunogenic peptide #489.  
XX  
KM Vaccine; virucide; antigen; autoimmune disease; infection;  
KM immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;  
KM breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;  
KM pancreas tumor; stomach tumor; bladder tumor; kidney tumor;  
KM hodgekin's lymphoma.  
XX  
OS Hepatitis B virus.  
XX  
PN WO2004108753-A1.  
XX  
PD 16-DEC-2004.  
XX  
PF 10-JUN-2004; 2004WO-AU000775.  
XX  
PR 10-JUN-2003; 2003AU-00902875.  
XX  
PR 25-MAR-2004; 2004AU-00901589.  
XX  
PA (UYME ) UNIV MELBOURNE.  
XX  
PI Kent SJ;  
XX  
DR WPI; 2005-031657/03.

XX Use of at least one set of peptides in the preparation of a medicament  
 PT for modulating an immune response, and for treating cancer or yeast,  
 PT viral, bacterial, protozoal and mycoplasma infections.

PS Disclosure; SEQ ID NO 2108; 645bp; English.

XX The invention relates to the use of at least one set of peptides in the  
 CC preparation of a medicament for modulating an immune response, where  
 CC individual peptides of a respective set comprise different portions of an  
 CC amino acid sequence corresponding to a single polypeptide of interest and  
 CC display partial sequence identity or similarity to at least one other  
 CC peptide of the same set of peptides (i.e. they are overlapping). Also  
 CC included are an antigen-presenting cell which has been contacted with the  
 CC peptides above and thus presents the peptides, a population of such  
 CC antigen-presenting cells, a process for producing antigen-presenting  
 CC cells for modulating an immune response to a polypeptide of interest, a  
 CC method for producing antigen-specific lymphocytes, a composition  
 CC comprising at least one set of the peptides (and a carrier and/or  
 CC diluent), a method for modulating an immune response to a polypeptide of  
 CC interest comprising administering to a patient in need at least one set  
 CC of the peptides, a method for treatment and/or prophylaxis of a disease  
 CC or condition associated with the presence of a polypeptide of interest  
 CC and a composition of matter for modulating an immune response in a  
 CC subject to a target antigen. The polypeptide of interest is also a  
 CC disease- or condition-associated polypeptide that is a polypeptide  
 CC produced by a pathogenic organism or a cancer, and produced by a  
 CC pathogenic organism selected from yeast, viruses, bacteria, helminths,  
 CC protozoans and mycoplasmas. The disease- or condition-associated  
 CC polypeptide is produced by a cancer selected from melanoma, lung cancer,  
 CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic  
 CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant  
 CC lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured  
 CC antigen-presenting cells or their precursors are useful in the  
 CC preparation of a medicament for the treatment of a disease or condition  
 CC in a subject, which disease or condition is associated with the presence  
 CC of aberrant expression of a target antigen, where the antigen-presenting  
 CC cells or their precursors have not been subjected to activating  
 CC conditions but have been contacted with an antigen that corresponds to  
 CC the target antigen to express a processed or modified form of the antigen  
 CC for presentation to the subject's immune system. The present sequence is  
 CC one of a set of overlapping immunogenic peptides derived from a Hepatitis  
 CC B virus protein.

XX Sequence 15 AA;

XX Query Match 60.0%; Score 18; DB 9; Length 15;  
 XX Best Local Similarity 25.0%; Pred. No. 2.4e+02;  
 XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 |  
 DB 4 CASSSSSC 11

RESULT 6  
 ADV23687

ID ADV23687 standard; peptide; 15 AA.

XX ADV23687;

DT 10-MAR-2005 (first entry)

DE HBV immunogenic peptide #488.

XX Vaccine; virucide; antigen; autoimmune disease; infection;  
 KM immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;  
 KM breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;  
 KM pancreas tumor; stomach tumor; bladder tumor; kidney tumor;  
 KM hodgkin's lymphoma.

XX Hepatitis B virus.

XX

PN WO2004108753-A1.

XX 16-DEC-2004.

XX 10-JUN-2004; 2004WO-AU000775.

XX 10-JUN-2003; 2003AU-00902875.

PR 25-MAR-2004; 2004AU-00901589.

XX (UYME ) UNIV MELBOURNE.

PI Kent St;

DR WPI; 2005-031657/03.

XX Use of at least one set of peptides in the preparation of a medicament  
 PT for modulating an immune response, and for treating cancer or yeast,  
 PT viral, bacterial, protozoal and mycoplasma infections.

PS Disclosure; SEQ ID NO 2107; 645bp; English.

XX The invention relates to the use of at least one set of peptides in the  
 CC preparation of a medicament for modulating an immune response, where  
 CC individual peptides of a respective set comprise different portions of an  
 CC amino acid sequence corresponding to a single polypeptide of interest and  
 CC display partial sequence identity or similarity to at least one other  
 CC peptide of the same set of peptides (i.e. they are overlapping). Also  
 CC included are an antigen-presenting cell which has been contacted with the  
 CC peptides above and thus presents the peptides, a population of such  
 CC antigen-presenting cells, a process for producing antigen-presenting  
 CC cells for modulating an immune response to a polypeptide of interest, a  
 CC method for producing antigen-specific lymphocytes, a composition  
 CC comprising at least one set of the peptides (and a carrier and/or  
 CC diluent), a method for modulating an immune response to a polypeptide of  
 CC interest comprising administering to a patient in need at least one set  
 CC of the peptides, a method for treatment and/or prophylaxis of a disease  
 CC or condition associated with the presence of a polypeptide of interest  
 CC and a composition of matter for modulating an immune response in a  
 CC subject to a target antigen. The polypeptide of interest is also a  
 CC disease- or condition-associated polypeptide that is a polypeptide  
 CC produced by a pathogenic organism or a cancer, and produced by a  
 CC pathogenic organism selected from yeast, viruses, bacteria, helminths,  
 CC protozoans and mycoplasmas. The disease- or condition-associated  
 CC polypeptide is produced by a cancer selected from melanoma, lung cancer,  
 CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic  
 CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant  
 CC lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured  
 CC antigen-presenting cells or their precursors are useful in the  
 CC preparation of a medicament for the treatment of a disease or condition  
 CC in a subject, which disease or condition is associated with the presence  
 CC of aberrant expression of a target antigen, where the antigen-presenting  
 CC cells or their precursors have not been subjected to activating  
 CC conditions but have been contacted with an antigen that corresponds to  
 CC the target antigen to express a processed or modified form of the antigen  
 CC for presentation to the subject's immune system. The present sequence is  
 CC one of a set of overlapping immunogenic peptides derived from a Hepatitis  
 CC B virus protein.

XX Sequence 15 AA;

XX Query Match 60.0%; Score 18; DB 9; Length 15;  
 XX Best Local Similarity 25.0%; Pred. No. 2.4e+02;  
 XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 |  
 DB 8 CASSSSSC 15

RESULT 7  
 ABB42273

ID ABB42273 standard; peptide; 24 AA.

XX

```

AC ABB42273;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #9779 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 34908; 639pp + Sequence Listing; English.
XX
SQ Sequence 24 AA;
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
QY 4 CXXXXXXC 11
DB 12 CTTSTTTC 19
XX
Query Match 60.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
RESULT 8
ID AAM36081 standard; protein; 24 AA.
XX
AC AAM36081;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #10118 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX

```

```

XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-48897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 36350; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see A131315-A157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 24 AA;
XX
QY 4 CXXXXXXC 11
DB 12 CTTSTTTC 19
XX
Query Match 60.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
RESULT 9
ID AAM75973 standard; protein; 24 AA.
XX
AC AAM75973;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36279.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

```

DR WPI; 2001-488900/53.  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human bone marrow.  
XX  
XX Example 4; SEQ ID NO 36279; 658bp + Sequence Listing; English.  
PS  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention  
XX  
SQ Sequence 24 AA;  
XX  
Query Match 60.0%; Score 18; DB 4; Length 24;  
Best Local Similarity 25.0%; Pred. No. 2.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
OY 4 CXXXXXXC 11  
DB 12 CTTSTTTC 19  
XX  
RESULT 10  
AAM63159 standard; protein; 24 AA.  
XX  
AC AAM63159;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35264.  
XX  
XX Human; brain expressed exon; gene expression analysis; probe; microarray;  
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000667.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
XX  
PR 26-MAY-2000; 2000US-0207456P.  
XX  
PR 30-JUN-2000; 2000US-00608408.  
XX  
PR 03-AUG-2000; 2000US-00632366.  
XX  
PR 21-SEP-2000; 2000US-0234687P.  
XX  
PR 27-SEP-2000; 2000US-0236359P.  
XX  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains.  
XX  
PS Example 4; SEQ ID NO 35264; 650bp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention  
XX

SQ Sequence 24 AA;  
XX  
Query Match 60.0%; Score 18; DB 4; Length 24;  
Best Local Similarity 25.0%; Pred. No. 2.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
OY 4 CXXXXXXC 11  
DB 12 CTTSTTTC 19  
XX  
RESULT 11  
ABG57702 standard; peptide; 24 AA.  
XX  
AC ABG57702;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human liver peptide, SEQ ID No 36350.  
XX  
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
KW hypercholesterolaemia; coronary heart disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200157273-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000664.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
XX  
PR 26-MAY-2000; 2000US-0207456P.  
XX  
PR 30-JUN-2000; 2000US-00608408.  
XX  
PR 03-AUG-2000; 2000US-00632366.  
XX  
PR 21-SEP-2000; 2000US-0234687P.  
XX  
PR 27-SEP-2000; 2000US-0236359P.  
XX  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488998/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human adult liver.  
XX  
XX  
PS Claim 27; SEQ ID NO 36350; 658bp; English.  
XX  
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult liver.  
CC (II) may be used for predicting, measuring and displaying gene expression  
CC in samples derived from human adult liver. The genes identified may be  
CC involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
CC associated with coronary heart disease. ABG47348-ABG59930 represent human  
CC liver single exon encoded peptides of the invention. Note: The sequence  
CC information for this patent does not appear in the printed specification  
CC but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 24 AA;  
XX  
Query Match 60.0%; Score 18; DB 4; Length 24;  
Best Local Similarity 25.0%; Pred. No. 2.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
OY 4 CXXXXXXC 11

Db 12 CTTSTTTC 19

RESULT 12  
ABB68695  
XX ABB68695 standard; protein; 69 AA.  
XX  
AC ABB68695;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 32877.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical.  
XX Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
XX 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX N-PSDB; ABL12798.  
DR WPI; 2001-656860/75.  
XX  
DR N-PSDB; ABL12798.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signaling and cell-cell  
XX interactions.  
XX  
PS Disclosure; SEQ ID NO 32877; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
XX ABB72072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 69 AA;  
XX  
XX Query Match 60.0%; Score 18; DB 4; Length 69;  
XX Best Local Similarity 25.0%; Pred. No. 3.3e+02;  
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
XX  
QY 4 CXXXXXXC 11  
XX |  
Db 9 CCGAASSC 16

RESULT 13  
ADP30715  
ID ADP30715 standard; protein; 69 AA.  
XX  
AC ADP30715;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1482.  
XX

KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411011P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 18-APR-2003; 2003US-0463700P.  
XX 18-APR-2003; 2003US-0463708P.  
XX 18-APR-2003; 2003US-0463716P.  
XX 18-APR-2003; 2003US-0463732P.  
XX 02-MAY-2003; 2003US-0467199P.  
XX 02-MAY-2003; 2003US-0467201P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 19-MAY-2003; 2003US-0471306P.  
XX 19-MAY-2003; 2003US-0471336P.  
XX 22-MAY-2003; 2003US-0472420P.  
XX 22-MAY-2003; 2003US-0472430P.  
XX 09-JUN-2003; 2003US-0476609P.  
XX 09-JUN-2003; 2003US-0476641P.  
XX 08-JUL-2003; 2003US-0485218P.  
XX 08-JUL-2003; 2003US-0485223P.  
XX 08-JUL-2003; 2003US-0485224P.  
XX 08-JUL-2003; 2003US-0485325P.  
XX 14-JUL-2003; 2003US-0486446P.  
XX 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halanbeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2713; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC anti-inflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 69 AA;

Query Match 60.0%; Score 18; DB 8; Length 69;  
Best Local Similarity 25.0%; Pred. No. 3.3e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 4 CXXXXXXC 11  
|  
DB 18 CTTTAAAC 25

RESULT 14  
ABB39058  
ID ABB39058 standard; peptide; 91 AA.  
XX  
XX ABB39058;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #6564 encoded by human foetal liver single exon probe.  
XX  
XX Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
XX Homo sapiens.  
XX  
XX WO200157277-A2.  
XX  
XX 09-AUG-2001.  
XX  
PD 30-JAN-2001; 2001WO-US000669.  
XX  
PF 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX

DR WPI; 2001-483447/52.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human fetal liver.  
XX  
XX Claim 27; SEQ ID NO 31693; 639pp + Sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for measuring  
CC human gene expression in a sample derived from human foetal liver. The  
CC single exon nucleic acid probes may be used for predicting, measuring and  
CC displaying gene expression in samples derived from human fetal liver. The  
CC present sequence is a peptide encoded by a single exon nucleic acid probe  
CC of the invention. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPo at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 91 AA;

Query Match 60.0%; Score 18; DB 4; Length 91;  
Best Local Similarity 25.0%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 4 CXXXXXXC 11  
|  
DB 29 CSSSSSSC 36

RESULT 15  
AAM32549  
ID AAM32549 standard; protein; 91 AA.  
XX  
XX AAM32549;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Peptide #6586 encoded by probe for measuring placental gene expression.  
XX  
XX Probe; microarray; human; placenta; antenatal diagnosis;  
XX genetic disorder.  
XX  
XX Homo sapiens.  
XX  
XX WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000663.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-488897/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human placenta.  
XX  
XX Claim 27; SEQ ID NO 32818; 654pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes (SENP;  
CC see AAI1315-AI157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders

XX  
SQ Sequence 91 AA;

Query Match 60.0%; Score 18; DB 4; Length 91;  
Best Local Similarity 25.0%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 CXXXXXXC 11  
|  
Db 29 CXXXXSSSC 36

RESULT 16  
AAM72290  
ID AAM72290 standard; protein; 91 AA.

AC AAM72290;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 32596.

XX Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukaemia; lymphoma; myeloma.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000668.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488900/53.

DR WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human bone marrow.

PS Example 4; SEQ ID NO 32596; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is a

CC protein encoded by one of the probes of the invention

XX SQ Sequence 91 AA;

Query Match 60.0%; Score 18; DB 4; Length 91;

Best Local Similarity 25.0%; Pred. No. 3.5e+02; Mismatches 6; Indels 0; Gaps 0;

OY 4 CXXXXXXC 11  
|  
Db 29 CXXXXSSSC 36

RESULT 17  
AAM59710  
ID AAM59710 standard; protein; 91 AA.

XX AAM59710;

AC 05-NOV-2001 (first entry)

DT Human brain expressed single exon probe encoded protein SEQ ID NO: 31815.

DE Human; brain expressed exon; gene expression analysis; probe; microarray;

XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000667.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-483446/52.

DR WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains.

PS Example 4; SEQ ID NO 31815; 650bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of

CC the probes of the invention

XX SQ Sequence 91 AA;

Query Match 60.0%; Score 18; DB 4; Length 91;

Best Local Similarity 25.0%; Pred. No. 3.5e+02; Mismatches 6; Indels 0; Gaps 0;

OY 4 CXXXXXXC 11  
|  
Db 29 CXXXXSSSC 36

RESULT 18  
ABG53976  
ID ABG53976 standard; peptide; 91 AA.

AC ABG53976;

DT 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID NO 32624.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

KW hypercholesterolaemia; coronary heart disease.

OS Homo sapiens.

PN WO200157273-A2.

```

PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000664.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488698/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
XX Claim 27; SEQ ID NO 32624; 658bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridizes at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABG47348-ABG5930 represent human
XX liver single exon encoded peptides of the invention. Note: The sequence
XX information for this patent does not appear in the printed specification
XX but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 91 AA;
XX
XX Query Match 60.0%; Score 18; DB 4; Length 91;
XX Best Local Similarity 25.0%; Pred. No. 3.5e+02;
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
OY 4 CXXXXXXC 11
XX
DB 29 CXXXXXXC 36
XX
RESULT 19
ABG42105
ID ABG42105 standard; peptide; 91 AA.
XX
XX ABG42105;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human peptide encoded by genome-derived single exon probe SEQ ID 31770.
XX
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Budlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX
XX Homo sapiens.
XX
XX WO200166003-A2.
XX

```

```

PD 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
XX Claim 27; SEQ ID NO 31770; 634bp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of probes
XX (I) the novel set of probes which hybridize at high stringency to a nucleic
XX acid expressed in the human lung; measuring gene expression in a sample
XX derived from human lung, comprising (a) contacting the array with a
XX collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of the
XX array; identifying exons in a eukaryotic genome, comprising (a)
XX algorithmically predicting at least one exon from genomic sequences of
XX the eukaryote; and (b) detecting specific hybridization of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridization to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene expression
XX analysis, and for identifying exons in a gene, particularly using human
XX lung derived mRNA and for the study of lung diseases such as asthma, lung
XX cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
XX disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
XX tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
XX Budlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
XX histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
XX Karsenger syndrome, fibrocystic pulmonary dysplasia, primary ciliary
XX dyskinesia, pulmonary hypertension and hyaline membrane disease. The
XX present sequence is a peptide/protein encoded by a single exon probe of
XX the invention. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 91 AA;
XX
XX Query Match 60.0%; Score 18; DB 5; Length 91;
XX Best Local Similarity 25.0%; Pred. No. 3.5e+02;
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
OY 4 CXXXXXXC 11
XX
DB 29 CXXXXXXC 36
XX

```





XX (GENO-) GENOME THERAPEUTICS CORP.  
PA Rubenfield MJ, Nolling J, DeLoughery C, Bush D;  
XX MPI: 2003-615309/58.  
DR N-PsDB; ABD16738.  
XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PR pathological conditions resulting from bacterial infection.  
XX  
PS Disclosure; SEQ ID NO 31913; 455bp; English.  
XX  
XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-*P. aeruginosa* drugs, as templates for recombinant  
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
CC of *Pseudomonas* species using biochip technology. Sequences ABO67826-  
CC ABO64396 represent *P. aeruginosa* polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)  
XX  
SQ Sequence 99 AA:  
Query Match 60.0%; Score 18; DB 7; Length 99;  
Best Local Similarity 25.0%; Pred. No. 3.6e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
OY 4 CXXXXXXC 11  
Db 24 CSSTASAC 31  
RESULT 22  
ADP30709  
ID ADP30709 standard; protein; 99 AA.  
XX  
AC ADP30709;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1476.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Halehan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JCP, Wu G, Zhang H;  
XX MPI: 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2707; 428bp; English.  
XX

CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.

XX  
SQ Sequence 99 AA;

Query Match 60.0%; Score 18; DB 8; Length 99;  
Best Local Similarity 25.0%; Pred. No. 3.6e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 77 CAAATATC 84

RESULT 23

ID ABO74878 standard; protein; 107 AA.

XX ABO74878;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #7053.

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US651795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

DR N-PSDB; ABD08449.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 23624; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using bioclip technology. Sequences ABO67826-

CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html

SQ Sequence 107 AA;

Query Match 60.0%; Score 18; DB 7; Length 107;  
Best Local Similarity 25.0%; Pred. No. 3.6e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 29 CSASSSAC 36

RESULT 24

ID ABO73394 standard; protein; 115 AA.

XX ABO73394;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #5569.

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US651795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI; 2003-615309/58.

DR N-PSDB; ABD06965.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 22140; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using bioclip technology. Sequences ABO67826-  
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html

SQ Sequence 115 AA;

Query Match 60.0%; Score 18; DB 7; Length 115;  
Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 83 CSAATSTC 90

RESULT 25  
ADP04777  
ID ADF04777 standard; protein; 122 AA.  
XX  
AC ADF04777;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Bacterial polypeptide #890.  
XX  
KM Proteus mirabilis infection; bacterial infection; antibacterial;  
KW immunostimulant.  
XX  
OS Proteus mirabilis.  
XX  
PN US6605709-B1.  
XX  
PD 12-AUG-2003.  
XX  
PF 05-APR-2000; 2000US-00543681.  
XX  
PR 09-APR-1999; 99US-0128706P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Breton GL;  
XX  
PI MPI: 2003-895291/82.  
DR N-PSDB; ADP00605.  
XX  
PT New Proteus mirabilis polypeptides and polynucleotides, useful as  
PT reagents for diagnosis of bacterial disease, as components of  
PT antibacterial vaccines, as targets for antibacterial drugs, or as  
PT biocontrol agents for plants.  
XX  
PS Disclosure; SEQ ID NO 5062; 870pp; English.  
XX  
CC The invention relates to new Proteus mirabilis polypeptides and  
CC polynucleotides. The invention also relates to antibodies against the  
CC polypeptides, methods for producing the polypeptides, a method of  
CC generating vaccines for immunising an individual against P. mirabilis, a  
CC method for evaluating a compound for the ability to bind a P. mirabilis  
CC polypeptide and a method for screening test compounds for anti-bacterial  
CC activity. The polypeptides and polynucleotides are useful as molecular  
CC targets for diagnosing, preventing and treating pathological conditions  
CC resulting from bacterial infection, as reagents for diagnosis of  
CC bacterial diseases, as components of antibacterial vaccines, as targets  
CC for antibacterial drugs or as bio-control agents for plants. This  
CC sequence represents a Proteus mirabilis polypeptide of the invention.  
XX  
SQ Sequence 122 AA;  
XX  
Query Match 60.0%; Score 18; DB 7; Length 122;  
Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
OY 4 CXXXXXXC 11  
Db 99 CAAASAC 106  
XX  
RESULT 26  
ADP30703  
ID ADP30703 standard; protein; 135 AA.  
XX  
AC ADP30703;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1470.  
XX

KM Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; Immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476099P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-048524P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halebeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
PS Claim 1; SEQ ID NO 2701; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 135 AA;

Query Match 60.0%; Score 18; DB 8; Length 135;  
Best Local Similarity 25.0%; Pred. No. 3.8e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 4 CXXXXXXC 11  
DB 105 CTTTATC 112

RESULT 27  
ABO75730  
ID ABO75730 standard; protein; 144 AA.  
XX  
AC ABO75730;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Pseudomonas aeruginosa polypeptide #7905.  
XX  
KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX  
OS Pseudomonas aeruginosa.  
XX  
PN US6551795-B1.  
XX  
PD 22-APR-2003.  
XX  
PF 18-FEB-1999; 99US-00252991.  
XX  
PR 18-FEB-1998; 98US-0074788P.  
PR 27-JUL-1998; 98US-0094190P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX  
DR WPI; 2003-615309/58.  
XX  
DR N-PSDB; ABD09301.  
XX  
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of

PT pathological conditions resulting from bacterial infection.  
XX  
XX Disclosure; SEQ ID NO 24476; 455bp; English.  
XX  
XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 144 AA;

Query Match 60.0%; Score 18; DB 7; Length 144;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 4 CXXXXXXC 11  
DB 36 CSATSNAC 43

RESULT 28  
ADP31453  
ID ADP31453 standard; protein; 144 AA.  
XX  
AC ADP31453;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2220.  
XX  
KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486860P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Halsehan L, Lannemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX MPI; 2004-348438/32.  
DR  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PI genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 3451; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antinflammatory, immunosuppressive, antibacterial and antiviral. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
XX Sequence 144 AA;

Query Match 60.0%; Score 18; DB 8; Length 144;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 4 CXXXXXXC 11  
Db 20 CTAATTAC 27  
RESULT 29  
ADX90450 standard; protein; 144 AA.  
XX  
XX ADX90450;  
XX  
XX 21-APR-2005 (first entry)  
XX  
XX Plant full length insert polypeptide seqid 53114.  
XX  
XX plant protectant; plant growth regulant; gene therapy; plant;  
XX recombinant DNA construct; physical array; plant breeding marker;  
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
XX extreme osmotic condition; pathogen tolerance; pest tolerance;  
XX growth rate; cell cycle pathway; disease resistance;  
XX galactomannan production; lignin production; plant growth regulator;  
XX yield; plant growth; plant development; seed oil; protein yield;  
XX  
XX Unidentified.  
XX  
XX US2004034888-A1.  
XX  
XX 19-FEB-2004.  
XX  
XX 28-APR-2003; 2003US-00425114.  
XX  
XX 06-MAY-1999; 98US-00304517.  
XX 05-NOV-2001; 2001US-00985678.  
XX  
XX (LIU/J) LIU J.  
XX (ZHOU/) ZHOU Y.  
XX (KOVA/) KOVALIC D K.  
XX (SCREA/) SCREAN S E.  
XX (TABAK/) TABASKA J E.  
XX (CAO/Y) CAO Y.  
XX  
XX Liu J, Zhou Y, Kovalic DK, Screen SR, Tabaska JE, Cao Y;  
XX  
XX MPI; 2004-180133/17.  
DR  
XX  
XX New recombinant DNA construct, useful for improving plant tolerance to  
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
PT pests, for conferring increased resistance to plant disease, or for  
PT improving yield.  
XX  
XX  
PS Claim 1; SEQ ID NO 53114; 15bp; English.  
XX  
XX The invention describes a recombinant DNA construct comprising a  
CC polynucleotide consisting of a sequence encoding an amino acid sequence  
CC available in electronic form from the US patent office at  
CC ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide  
CC of the invention are also useful in physical arrays of molecules and as  
CC plant breeding markers. The recombinant DNA construct is useful for  
CC improving plant tolerance to cold, heat, drought, herbicides, extreme  
CC osmotic conditions, pathogens or pests, for manipulating growth rate in  
CC plant cells by modification of the cell cycle pathway, for conferring  
CC increased resistance to plant disease, for producing galactomannan,  
CC lignin or plant growth regulators, for increasing the rate of homologous  
CC recombination in plants, for improving yield by modification of  
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake  
CC or by providing improved plant growth and development under at least one  
CC stress condition or for modifying seed oil or protein yield and/or

CC content. This is the amino acid sequence of a plant full length insert  
 CC polypeptide that can be used in the recombinant DNA construct of the  
 CC invention.

XX  
 SQ Sequence 144 AA;

Query Match 60.0%; Score 18; DB 8; Length 144;  
 Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 4 CTSSSAAAC 11

RESULT 30

ID ABB45351 standard; protein; 150 AA.

XX ABB45351;

XX 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #30878.

KM Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Salmonella paratyphi.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362599P.

XX (ELITRA) PHARM INC.

XX Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA49221.

XX Claim 25; SEQ ID NO 73275; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation; (7) identifying a gene in an operon required for  
 CC the gene product or that has an activity against a biological pathway;  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strings is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pat\_sequences

XX Sequence 150 AA;

Query Match 60.0%; Score 18; DB 6; Length 150;  
 Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 16 CATTASSC 23

RESULT 31

ID AAB41043 standard; protein; 153 AA.

XX AAB41043;

XX 08-FEB-2001 (first entry)

DE Human ORF807 polypeptide sequence SEQ ID NO:1614.

XX Human; open reading frame; ORF; detection; cytostatic; hepatotropic;

XX vulnary; antiproliferative; antiparkinsonian; neuroprotective;

XX anticonvulsant; osteopathic; antichratic; immunosuppressant; cardiac;

XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;

XX hypotensive; dermatological; immunosuppressive; antiinflammatory;

XX antiviral; antitubercular; antitumor; antineuritic; antihypertensive;

XX antineuritic; gene therapy; cancer; proliferative disorder; hypertension;

XX neurodegenerative disorder; osteoarthritis; graft vs host disease;

XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;

XX cholesterol ester storage; systemic lupus erythematosus; infection;

XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;

XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;

XX bone damage; cartilage damage; antiinflammatory disease; coagulation;

XX thrombosis; contraceptive.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008621.

XX 31-MAR-1999; 99US-0127607P.

XX 02-APR-1999; 99US-0127636P.

XX 30-MAR-2000; 2000US-00540763.

CC (CURA-) CURAGEN CORP.  
 CC Shinketsu RA, Leach M;  
 CC WPI; 2000-602362/57.  
 CC N-PSDB; AACT5252.  
 CC Novel nucleic acids and peptides derived from open reading frame X,  
 CC useful for treating e.g. cancers, proliferative disorders,  
 CC neurodegenerative disorders and cardiovascular disease.

XX Claim 11; Page 1302; 5507pp; English.  
 PS  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytosolic; hepatotropic; vulnery;  
 CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective; osteopathic;  
 CC antipsoriatic; antiparkinsonian; neurotropic; hepatotropic; vulnery;  
 CC anticonvulsant; antiallergic; immunosuppressant; immunostimulant;  
 CC cardiatic; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
 CC antiviral; antifungal; antineumatic; antihypertoid; and antiaemic. The  
 CC sequences can be used for determining the presence of or predisposition  
 CC to, or preventing or treating pathological conditions associated with an  
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
 CC used to treat cancers, proliferative disorders, neurodegenerative  
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
 XX  
 XX Sequence 153 AA;  
 SQ  
 Query Match 60.0%; Score 18; DB 3; Length 153;  
 Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 4 CXXXXXXC 11  
 Db 78 CSSSASC 85  
 RESULT 32  
 ID ABB09251 standard; protein; 153 AA.  
 XX  
 AC ABB09251;  
 XX  
 DT 25-JUN-2002 (first entry)  
 XX  
 DE Human ORFX protein sequence SEQ ID NO:18484.  
 XX  
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
 KW hyperproliferative disorder; psoriasis; benign tumor; haemorrhage;  
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
 KW hypertension; hypothyroidism; cholesterol ester storage disease;  
 KW immune deficiency; immune disorder; infectious disease;  
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
 KW myasthenia gravis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200192523-A2.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 29-MAY-2001; 2001WO-US010836.  
 XX  
 PR 30-MAY-2000; 2000US-0206132P.  
 PR 29-AUG-2000; 2000US-0228716P.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shimkets RA, Leach MD;  
 XX  
 DR WPI; 2002-106308/14.  
 DR N-PSDB; ABB25003.  
 XX  
 PT Novel human polypeptides and polynucleotides useful for diagnosing,

PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and autoimmune disorders.  
 XX  
 XX Disclosure; SEQ ID NO 18484; 1037pp; English.  
 XX  
 CC The present invention describes substantially purified human proteins  
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
 CC in the specification). ABB15762 to ABB17552 encode the human ORFX  
 CC proteins given in ABB00010 to ABB11500. ORFX proteins are useful for  
 CC treating or preventing a pathology associated with an ORFX-associated  
 CC disorder in humans, and in the manufacture of a medicament for treating a  
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
 CC sequences can be used in gene therapy. ORFX sequences can be used in the  
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
 CC storage disease, various immune deficiencies and disorders, infectious  
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 CC bone degenerative disorders, or periodontal disease, and for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues and conditions resulting from  
 CC systemic cytokine damage. N.B. The sequence data for this patent did not  
 CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 153 AA;  
 SQ  
 Query Match 60.0%; Score 18; DB 5; Length 153;  
 Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 4 CXXXXXXC 11  
 Db 78 CSSSASC 85  
 RESULT 33  
 ID ABO74421 standard; protein; 154 AA.  
 XX  
 AC ABO74421;  
 XX  
 DT 29-JUL-2004 (first entry)  
 XX  
 DE Pseudomonas aeruginosa polypeptide #6596.  
 XX  
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
 XX  
 OS Pseudomonas aeruginosa.  
 XX  
 PN US6551795-B1.  
 XX  
 PD 22-APR-2003.  
 XX  
 PF 18-FEB-1999; 99US-00252991.  
 XX  
 PR 18-FEB-1998; 98US-0074788P.  
 PR 27-JUL-1998; 98US-0094190P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
 XX  
 DR WPI; 2003-615309/58.  
 DR N-PSDB; ABO7992.  
 XX  
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of



PT pathological conditions resulting from bacterial infection.  
XX  
PS Disclosure; SEQ ID NO 23167; 455bp; English.  
XX  
XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-*P. aeruginosa* drugs, as templates for recombinant  
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-  
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 154 AA;  
  
Query Match 60.0%; Score 18; DB 7; Length 154;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
QY 4 CXXXXXXC 11  
DB 120 CAAASAC 127  
  
RESULT 34  
AB077836  
ID AB077836 standard; protein; 156 AA.  
XX  
AC AB077836;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE *Pseudomonas aeruginosa* polypeptide #10011.  
XX  
KM Bacterial infection; *Pseudomonas aeruginosa* infection; antibacterial.  
XX  
OS *Pseudomonas aeruginosa*.  
XX  
PN US651795-B1.  
XX  
PD 22-APR-2003.  
XX  
PF 18-FEB-1999; 99US-00252991.  
XX  
PR 18-FEB-1998; 98US-0074788P.  
XX  
PR 27-JUL-1998; 98US-0094190P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX  
DR WPI; 2003-615309/58.  
XX  
DR N-PSDB; ABD11407.  
XX  
PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
XX Disclosure; SEQ ID NO 26582; 455bp; English.  
XX  
CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,

CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-*P. aeruginosa* drugs, as templates for recombinant  
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-  
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 156 AA;  
  
Query Match 60.0%; Score 18; DB 7; Length 156;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
QY 4 CXXXXXXC 11  
DB 40 CSASSAAC 47  
  
RESULT 35  
ADP30761  
ID ADP30761 standard; protein; 159 AA.  
XX  
AC ADP30761;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1528.  
XX  
KM Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; Inflammatory; Immune; human secreted protein.  
XX  
OS *Homo sapiens*.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
XX  
PR 29-AUG-2002; 2002US-0406579P.  
XX  
PR 29-AUG-2002; 2002US-0406585P.  
XX  
PR 29-AUG-2002; 2002US-0406588P.  
XX  
PR 29-AUG-2002; 2002US-0406608P.  
XX  
PR 29-AUG-2002; 2002US-0406611P.  
XX  
PR 29-AUG-2002; 2002US-0406612P.  
XX  
PR 29-AUG-2002; 2002US-0406616P.  
XX  
PR 29-AUG-2002; 2002US-0406640P.  
XX  
PR 29-AUG-2002; 2002US-0406642P.  
XX  
PR 29-AUG-2002; 2002US-0406646P.  
XX  
PR 29-AUG-2002; 2002US-0406653P.  
XX  
PR 29-AUG-2002; 2002US-0406655P.  
XX  
PR 29-AUG-2002; 2002US-0406666P.  
XX  
PR 17-SEP-2002; 2002US-0410946P.  
XX  
PR 17-SEP-2002; 2002US-0410947P.  
XX  
PR 17-SEP-2002; 2002US-0410948P.  
XX  
PR 17-SEP-2002; 2002US-0410949P.  
XX  
PR 17-SEP-2002; 2002US-0410953P.  
XX  
PR 17-SEP-2002; 2002US-0410957P.  
XX  
PR 17-SEP-2002; 2002US-0410958P.  
XX  
PR 17-SEP-2002; 2002US-0410959P.  
XX  
PR 17-SEP-2002; 2002US-0410960P.  
XX  
PR 17-SEP-2002; 2002US-0410961P.  
XX  
PR 17-SEP-2002; 2002US-0410962P.  
XX  
PR 17-SEP-2002; 2002US-0411019P.  
XX  
PR 17-SEP-2002; 2002US-0411022P.  
XX  
PR 17-SEP-2002; 2002US-0411023P.  
XX  
PR 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411010P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485252P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haishan U, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 2759; 428bp; English.  
XX  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
XX  
SQ Sequence 159 AA;

Query Match 60.0%; Score 18; DB 8; Length 159;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 125 CAATATTC 132

RESULT 36  
ID ADP30758  
ADP30758 standard; protein; 162 AA.  
XX  
XX  
AC ADP30758;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1525.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0405588P.  
PR 29-AUG-2002; 2002US-040608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Hsiehan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 2756; 428bp; English.  
XX  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
XX  
SQ Sequence 162 AA;  
XX  
XX  
Query Match 60.0%; Score 18; DB 8; Length 162;  
Best Local Similarity 25.0%; Pred. NO. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
OY 4 CXXXXXXC 11  
DB 53 CAATTATC 60  
XX  
XX  
RESULT 37  
ADP30756  
ID ADP30756 standard; protein; 162 AA.  
XX  
XX  
AC ADP30756;  
XX  
XX  
DT 12-AUG-2004 (first entry)  
XX  
XX  
DE Human secreted protein SEQ ID #1523.  
XX  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW Cancer; Inflammatory; Immune; human secreted protein.  
XX  
XX  
OS Homo sapiens.  
XX  
XX  
PN MO2004035732-AZ.  
XX  
XX  
PD 29-APR-2004.  
XX  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX

PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 09-JUN-2003; 2003US-0476430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Hsiehan L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PI genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2754; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 162 AA;  
  
Query Match 60.0%; Score 18; DB 8; Length 162;  
Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
QY 4 CXXXXXXC 11  
DB 53 CAATTATTC 60  
  
RESULT 38  
ADP30757  
ID ADP30757 standard; protein; 162 AA.  
XX  
XX ADP30757;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1524.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
XX  
PD 28-AUG-2003; 2003WO-US026780.  
XX  
PF 29-AUG-2002; 2002US-0406576P.  
XX  
XX 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406656P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493572P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Haislan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2755; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 162 AA;  
  
Query Match 60.0%; Score 18; DB 8; Length 162;

Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 53 CATTATTC 60

## RESULT 39

ADP30813  
ID ADP30813 standard; protein; 162 AA.

AC ADP30813;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1580.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX WO2004035732-A2.

XX 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

XX 17-SEP-2002; 2002US-0411046P.

XX 17-SEP-2002; 2002US-0411048P.

XX 17-SEP-2002; 2002US-0411052P.

XX 17-SEP-2002; 2002US-0411055P.

XX 17-SEP-2002; 2002US-0411073P.

XX 17-SEP-2002; 2002US-0411082P.

XX 17-SEP-2002; 2002US-0411101P.

XX 17-SEP-2002; 2002US-0411111P.

XX 18-APR-2003; 2003US-0463700P.

XX 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu X, Lee E, Hestir K, Beaurang PA, Behrens D;

PI Halenbeck RF, Huang MM, Kochakota S, Halenhan L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

XX WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases

PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,

XX genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 2811; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule

CC encoding a polypeptide which is believed to be cytostatic,

CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is

CC available on WIPOMEB and is not in the specification.

XX SQ Sequence 162 AA;

QY Query Match 60.0%; Score 18; DB 8; Length 162;

Best Local Similarity 25.0%; Pred. No. 4e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

DB 4 CXXXXXXC 11

DB 107 CATTTC 114

## RESULT 40

ADP31269

ID ADP31269 standard; protein; 165 AA.

AC ADP31269;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2036.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX WO2004035732-A2.  
PN  
XX  
PD 29-APR-2004.  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411010P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-046720P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 09-JUN-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476509P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485234P.  
PR 08-JUL-2003; 2003US-0485335P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493572P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linnemann T;  
PI Pierce K, Wang Y, Wong JG, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3267; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 165 AA;  
XX  
Query Match 60.0%; Score 18; DB 8; Length 165;  
Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 4 CXXXXXXC 11  
Db 62 CTAATTTTC 69  
RESULT 41  
ADP31099  
ID ADP31099 standard; protein; 168 AA.  
XX  
XX ADP31099;  
AC  
XX  
DT 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #1866.  
DE  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
PN  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0463719P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Hakenbeck RF, Huang MM, Kothakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3097; 428pp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC

Query Match	Best Local Similarity	Score 18;	DB 8;	Length 168;
Matches 2;	Conservative	0;	Mismatches 6;	Indels 0;
				Gaps 0;
OY	4 CXXXXXXC 11			
DB	25 CATTTC 32			
RESULT 42				
ID	ADP30794			
AC	ADP30794 standard; protein; 171 AA.			
DT	ADP30794;			
XX	12-AUG-2004 (first entry)			
DE	Human secreted protein SEQ ID #1561.			
XX				
XX	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;			
XX	Cancer; Inflammatory; Immune; human secreted protein.			
OS	Homo sapiens.			
XX				
XX	WO2004035732-A2.			
XX				
XX	29-APR-2004.			
PD				
PF	28-AUG-2003; 2003WO-US026780.			
XX				
XX	29-AUG-2002; 2002US-0406576P.			
XX	29-AUG-2002; 2002US-0406579P.			
XX	29-AUG-2002; 2002US-0406585P.			
XX	29-AUG-2002; 2002US-0406588P.			
XX	29-AUG-2002; 2002US-0406608P.			
XX	29-AUG-2002; 2002US-0406611P.			
XX	29-AUG-2002; 2002US-0406612P.			
XX	29-AUG-2002; 2002US-0406616P.			
XX	29-AUG-2002; 2002US-0406640P.			
XX	29-AUG-2002; 2002US-0406642P.			
XX	29-AUG-2002; 2002US-0406646P.			
XX	29-AUG-2002; 2002US-0406653P.			
XX	29-AUG-2002; 2002US-0406655P.			
XX	29-AUG-2002; 2002US-0406666P.			
XX	17-SEP-2002; 2002US-0410946P.			
XX	17-SEP-2002; 2002US-0410947P.			
XX	17-SEP-2002; 2002US-0410948P.			
XX	17-SEP-2002; 2002US-0410949P.			
XX	17-SEP-2002; 2002US-0410953P.			
XX	17-SEP-2002; 2002US-0410957P.			
XX	17-SEP-2002; 2002US-0410958P.			
XX	17-SEP-2002; 2002US-0410959P.			
XX	17-SEP-2002; 2002US-0410960P.			
XX	17-SEP-2002; 2002US-0410961P.			
XX	17-SEP-2002; 2002US-0410962P.			
XX	17-SEP-2002; 2002US-0411019P.			
XX	17-SEP-2002; 2002US-0411022P.			
XX	17-SEP-2002; 2002US-0411023P.			
XX	17-SEP-2002; 2002US-0411024P.			
XX	17-SEP-2002; 2002US-0411032P.			
XX	17-SEP-2002; 2002US-0411035P.			
XX	17-SEP-2002; 2002US-0411037P.			
XX	17-SEP-2002; 2002US-0411041P.			
XX	17-SEP-2002; 2002US-0411045P.			
XX	17-SEP-2002; 2002US-0411046P.			
XX	17-SEP-2002; 2002US-0411048P.			
XX	17-SEP-2002; 2002US-0411052P.			

PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463722P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486466P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-049341P.  
PR 08-AUG-2003; 2003US-049370P.  
PR 08-AUG-2003; 2003US-049373P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
(FIVE-) FIVE PRIME THERAPEUTICS INC.  
PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
XX Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2792; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 171 AA;

Query Match 60.0%; Score 18; DB 8; Length 171;  
Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
Db 161 CATTATTC 168

RESULT 43  
ADP30793  
ID ADP30793 standard; protein; 171 AA.  
XX  
AC ADP30793;  
XX  
DT 12-AUG-2004 (first entry)

XX  
DE Human secreted protein SEQ ID #1560.  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX Homo sapiens.  
OS  
XX WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.



PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486448P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2791; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
XX Sequence 171 AA:  
SQ  
Query Match 60.0%; Score 18; DB 8; Length 171;  
Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 4 CXXXXXXC 11  
Db 161 CATTATTC 168  
RESULT 44  
AA010193  
ID AAG10193 standard; protein; 177 AA.  
XX  
XX AAG10193;  
AC  
XX  
XX 17-OCT-2000 (first entry)  
DT  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 8414.  
DE  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
XX EP1033405-A2.  
PN  
XX  
XX 06-SEP-2000.  
PD  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
PF  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-01444532P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-015658P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157177P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.

PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 60.0%; Score 18; DB 3; Length 177;  
Best Local Similarity 25.0%; Pred. No. 46+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 2; Conservative 0;

QY 4 CXXXXXC 11  
Db 87 CXXXXXC 94

RESULT 45  
ADP30755  
ID ADP30755 standard; protein; 177 AA.  
XX  
AC ADP30755;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1522.  
XX  
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halembeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
PI Pierce Y, Wang Y, Mong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2753; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC Immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 177 AA;  
Query Match 60.0%; Score 18; DB 8; Length 177;  
Best Local Similarity 25.0%; Pred. No. 46+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 4 CXXXXXXC 11  
Db 145 CTATATTC 152  
RESULT 46  
ADP31682  
ID ADP31682 standard; protein; 177 AA.  
XX  
AC ADP31682;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2449.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN MO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PE 28-AUG-2003; 2003MO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411039P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-048690P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

## (FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3680; 428pp; English.

XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
XX available on WIPOMEB and is not in the specification.

XX Sequence 177 AA;

XX Query Match 60.0%; Score 18; DB 8; Length 177;  
XX Best Local Similarity 25.0%; Pred. No. 4e+02;  
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
Db 39 CAAAAATC 46

ADTS6379  
ID ADTS6379 standard; protein; 177 AA.  
XX ADTS6379;  
XX ADTS6379;  
DT 13-JAN-2005 (first entry)

XX Plant polypeptide, SEQ ID 6456.

XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;  
XX disease resistance; galactomannan production; plant growth regulator;  
XX heat tolerance; herbicide tolerance; lignin production;  
XX extreme osmotic condition tolerance; pathogens resistance;  
XX pest resistance; yield improvement; seed oil yield; seed protein yield.

XX Viridiplantae.

XX US2004216190-A1.

XX 28-OCT-2004.

XX 18-DEC-2003; 2003US-00739930.

XX 28-APR-2003; 2003US-00424599.

XX 28-APR-2003; 2003US-00425115.

XX (KOVA/) KOVALIC D K.

XX Kovalic DK;

XX WPI; 2004-757369/74.

XX New recombinant DNA constructs useful in the field of biochemistry and  
PT genetics, and in particular for producing transgenic plants with improved  
PT biological characteristics.

XX Claim 2; SEQ ID NO 6456; 14pp; English.

XX The invention relates a recombinant DNA construct comprising a  
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:  
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences  
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,  
CC Arabidopsis, wheat and rape but the specification does not indicate which  
CC sequences is derived from which organism. Also included is a method of  
CC producing a plant having an improved property, comprising transforming a  
CC plant with a recombinant DNA construct comprising a promoter region  
CC functional in a plant cell operably joined to a polynucleotide encoding a  
CC polypeptide associated with the property, and growing the transformed  
CC plant. The property is selected from improving plant cold tolerance, for  
CC manipulating growth rate in plant cells by modification of the cell cycle  
CC pathway, for improving plant drought tolerance, for providing increased  
CC resistance to plant disease, for galactomannan production, for production  
CC of plant growth regulators, for improving plant heat tolerance, for  
CC improving plant tolerance to herbicides, for increasing the rate of  
CC homologous recombination in plants, for lignin production, for improving  
CC plant tolerance to extreme osmotic conditions, for improving plant  
CC tolerance to pathogens or pests, for yield improvement by modification of  
CC photosynthesis, for modifying seed oil yield and/or content, for  
CC modifying seed protein yield and/or content, for yield improvement by  
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake  
CC and for yield improvement by providing improved plant growth and  
CC development under at least one stress condition. The polynucleotide may  
CC also encode a plant transcription factor. The methods and compositions of  
CC the present invention are useful in the field of biochemistry and  
CC genetics, in particular for producing transgenic plants with improved  
CC biological characteristics such as increased yield, improved nitrogen  
CC flow, increasing plant tolerance to cold or heat, improving plant  
CC tolerance to extreme osmotic and drought conditions, and improving plant  
CC tolerance to plant pests or pathogens. They can also be used in physical  
CC arrays of molecules, plant breeding markers, computer-based storage and  
CC analysis systems. The present sequence is one of the 5544 plant protein  
CC sequences of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from USPTO at  
CC segdata.uspto.gov/sequence.html?docID=20040216190.

XX Sequence 177 AA;

XX Query Match 60.0%; Score 18; DB 8; Length 177;

Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 87 CXXXXXXC 94

## RESULT 48

ABM86593  
ID ABM86593 standard; protein; 182 AA.

XX ABM86593;

DT 02-JUN-2005 (first entry)

DE Rice abiotic stress responsive polypeptide SEQ ID NO:4839.

KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.

OS Oryza sativa.

PN WO2003008540-A2.

PD 30-JAN-2003.

PF 21-JUN-2002; 2002WO-US019668.

PR 22-JUN-2001; 2001US-0300112P.

PR 24-AUG-2001; 2001US-0314662P.

PR 26-SEP-2001; 2001US-0325277P.

PR 21-NOV-2001; 2001US-0332132P.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,

PI Moughamer T, Provart N, Rieke D, Zhu T;

XX MPI; 2003-248011/24.

PS Claim 1; SEQ ID NO 4839; 89pp; English.

XX The invention relates to novel abiotic stress responsive polynucleotides  
CC and polypeptides. Also disclosed are vectors, expression cassettes, host  
CC cells, and plants containing such polynucleotides. Also disclosed are  
CC methods for using the polynucleotides and polypeptides to alter the  
CC responsiveness of a plant to abiotic stress. The invention is useful in  
CC agriculture. The nucleic acid is useful for determining whether a test  
CC plant has been exposed to an abiotic stress condition. It is also useful  
CC for selecting an agent that alters abiotic stress regulated  
CC polynucleotide expression in a plant cell, and to identify a homolog or  
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid  
CC molecule and the polypeptide encoded by it are useful in altering the  
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt  
CC stress, osmotic stress or any of their combinations. The present sequence  
CC is used in the exemplification of the invention

XX Sequence 182 AA;

Query Match 60.0%; Score 18; DB 7; Length 182;

Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 15 CSTAGAC 22

## RESULT 49

ABM89559  
ID ABM89559 standard; protein; 185 AA.

XX ABM89559;

DT 02-JUN-2005 (first entry)

DE Rice abiotic stress responsive polypeptide SEQ ID NO:7805.

KW abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.

OS Oryza sativa.

PN WO2003008540-A2.

PD 30-JAN-2003.

PF 21-JUN-2002; 2002WO-US019668.

PR 22-JUN-2001; 2001US-0300112P.

PR 24-AUG-2001; 2001US-0314662P.

PR 26-SEP-2001; 2001US-0325277P.

PR 21-NOV-2001; 2001US-0332132P.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Kreps J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,

PI Moughamer T, Provart N, Rieke D, Zhu T;

XX MPI; 2003-248011/24.

PS Claim 1; SEQ ID NO 7805; 89pp; English.

XX The invention relates to novel abiotic stress responsive polynucleotides  
CC and polypeptides. Also disclosed are vectors, expression cassettes, host  
CC cells, and plants containing such polynucleotides. Also disclosed are  
CC methods for using the polynucleotides and polypeptides to alter the  
CC responsiveness of a plant to abiotic stress. The invention is useful in  
CC agriculture. The nucleic acid is useful for determining whether a test  
CC plant has been exposed to an abiotic stress condition. It is also useful  
CC for selecting an agent that alters abiotic stress regulated  
CC polynucleotide expression in a plant cell, and to identify a homolog or  
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid  
CC molecule and the polypeptide encoded by it are useful in altering the  
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt  
CC stress, osmotic stress or any of their combinations. The present sequence  
CC is used in the exemplification of the invention

XX Sequence 185 AA;

Query Match 60.0%; Score 18; DB 7; Length 185;

Best Local Similarity 25.0%; Pred. No. 4.1e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 148 CSTATTSC 155

## RESULT 50

ADP31576  
ID ADP31576 standard; protein; 192 AA.

XX ADP31576;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2343.



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 14.487 Seconds  
(without alignments)  
92.983 Million cell updates/sec

Title: US-09-932-322-2

Sequence: 1 XXXXXXXXXXXXX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 1000 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	60.0	177	2	F96719
2	18	60.0	355	2	A34597
3	18	60.0	367	2	C39590
4	18	60.0	492	2	S49147
5	18	60.0	843	1	JDVLJ1
6	18	60.0	843	1	JDVLJ2
7	18	60.0	1743	2	T26859
8	18	60.0	2543	2	T31687
9	18	60.0	2543	2	T31687
10	17	56.7	89	2	A44036
11	17	56.7	86	2	S72742
12	17	56.7	101	2	T42262
13	17	56.7	120	2	S72979
14	17	56.7	122	2	TJ0150
15	17	56.7	123	2	S25574
16	17	56.7	135	1	RWHUVY
17	17	56.7	139	2	A71123
18	17	56.7	145	2	H87436
19	17	56.7	182	2	T22305
20	17	56.7	203	2	E70641
21	17	56.7	212	2	H75277
22	17	56.7	224	2	A41128
23	17	56.7	230	2	S09778
24	17	56.7	238	1	JH0367
25	17	56.7	243	2	A35871
26	17	56.7	248	2	T47902
27	17	56.7	267	2	E87269
28	17	56.7	271	2	F70771
29	17	56.7	272	1	S72790

30	17	56.7	284	2	IS1172
31	17	56.7	292	2	C64795
32	17	56.7	292	2	D90710
33	17	56.7	292	2	H85560
34	17	56.7	297	2	A98294
35	17	56.7	297	2	AH2989
36	17	56.7	305	2	AE0857
37	17	56.7	305	2	G65056
38	17	56.7	307	2	B85925
39	17	56.7	307	2	A91080
40	17	56.7	309	2	T32376
41	17	56.7	347	2	B53522
42	17	56.7	348	2	T36890
43	17	56.7	350	2	C84548
44	17	56.7	352	2	B87507
45	17	56.7	365	2	G71287
46	17	56.7	376	2	D97175
47	17	56.7	400	2	C34443
48	17	56.7	400	2	AE1988
49	17	56.7	408	1	WMNVIA
50	17	56.7	408	2	A72869
51	17	56.7	412	2	AG3637
52	17	56.7	419	2	C40901
53	17	56.7	429	2	T16856
54	17	56.7	439	2	A36385
55	17	56.7	456	2	S20597
56	17	56.7	461	2	A31858
57	17	56.7	467	2	JC6150
58	17	56.7	471	2	A35480
59	17	56.7	471	2	A18446
60	17	56.7	489	2	S71955
61	17	56.7	492	1	WMAD55
62	17	56.7	539	2	H84640
63	17	56.7	541	2	T15299
64	17	56.7	544	2	S44814
65	17	56.7	547	1	SS1475
66	17	56.7	567	2	T49942
67	17	56.7	577	2	B37057
68	17	56.7	577	2	E86212
69	17	56.7	615	2	T00925
70	17	56.7	631	2	T00925
71	17	56.7	638	2	D86477
72	17	56.7	655	2	A59430
73	17	56.7	658	2	T19487
74	17	56.7	662	2	T17211
75	17	56.7	667	2	A48579
76	17	56.7	687	1	PFPMF
77	17	56.7	788	2	A37057
78	17	56.7	843	1	JDVLJ3
79	17	56.7	853	1	H70939
80	17	56.7	934	1	A34372
81	17	56.7	1028	2	T34360
82	17	56.7	1156	2	T23308
83	17	56.7	1279	2	A47363
84	17	56.7	1287	2	I46032
85	17	56.7	1345	2	S55669
86	17	56.7	1474	2	D88550
87	17	56.7	1522	2	H88380
88	17	56.7	1747	2	A45974
89	17	56.7	1857	2	S31212
90	17	56.7	1888	2	S78476
91	17	56.7	2395	1	S50820
92	17	56.7	2533	2	T28675
93	17	56.7	2533	2	T28674
94	17	56.7	2844	2	S28291
95	17	56.7	3124	2	A40020
96	17	56.7	3229	2	S27852
97	16	53.3	38	2	A47307
98	16	53.3	39	2	B41933
99	16	53.3	44	2	S38275
100	16	53.3	45	1	PFYZ
101	16	53.3	55	1	NTNMB2
102	16	53.3	57	2	E82522
					A46654

transcription fact  
c1g protein - Esc  
2-(5'-triphosphor  
2-(5'-triphosphor  
hypothetical prote  
permease [imported  
conserved hypothet  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
20k cytoplasmic  
probable iron-side  
hypothetical prote  
conserved hypothet  
conserved hypothet  
nifs family enzyme  
nitrogenase cofact  
nitrogenase cofact  
immediate-early pr  
early gene transac  
cis,cis-muconate t  
hypothetical prote  
hypothetical prote  
surface antigen se  
ribonuclease inhib  
ribonuclease-angio  
Gram-negative bact  
variant surface gl  
variant surface gl  
MDM2-like p53-bind  
early B1B 55k prot  
probable Rieske ir  
hypothetical prote  
F489.3 protein -  
cytochrome P450 cy  
hypothetical prote  
integrin beta-6 ch  
hypothetical prote  
hypothetical prote  
white protein - fr  
integrin beta-6 ch  
DNA-directed DNA p  
probable nitB prot  
complement C6 prec  
hypothetical prote  
hypothetical prote  
RNA helicase A - h  
nuclear DNA helica  
legument protein 7  
protein ZC84.6 [im  
protein T22F7.3 [i  
collagen alpha 1(X  
collagen alpha 1(X  
collagen alpha 1(X  
surface protein ty  
alpha-51D immobill  
alpha-51D-immobill  
hypothetical prote  
collagen alpha 1(X  
probable cell-surf  
heat-stable entero  
mating pheromone B  
toxin P1TX-II - sp  
furocin C - Myxoc  
neurotoxin B-II -  
hypothetical prote  
growth modulatory

103	16	53.3	57	2	B4654	growth modulatory
104	16	53.3	57	2	C4654	growth modulatory
105	16	53.3	62	2	S54356	metallothionein-2a
106	16	53.3	62	2	S54356	metallothionein-2c
107	16	53.3	72	1	WMB2P	ogr protein - phag
108	16	53.3	73	2	D82519	hypothetical prote
109	16	53.3	79	2	G82811	hypothetical prote
110	16	53.3	83	2	S11843	virB1 protein - Ag
111	16	53.3	83	2	AB3251	virA/G regulated p
112	16	53.3	91	2	G90824	epithelial cell g1
113	16	53.3	92	2	D37057	probable virion co
114	16	53.3	92	2	T30632	hypothetical prote
115	16	53.3	100	2	A13337	hypothetical prote
116	16	53.3	104	2	B81045	hypothetical prote
117	16	53.3	109	2	S24253	Ig heavy chain V r
118	16	53.3	114	2	H71018	hypothetical prote
119	16	53.3	115	2	B96664	probable RING zinc
120	16	53.3	118	1	PSSNK1	phospholipase A2 (
121	16	53.3	118	1	PSSNK3	phospholipase A2 (
122	16	53.3	118	2	D34860	phospholipase A2 (
123	16	53.3	118	2	E34860	phospholipase A2 (
124	16	53.3	118	2	G34860	phospholipase A2 (
125	16	53.3	118	2	F34860	phospholipase A2 (
126	16	53.3	118	2	H34860	phospholipase A2 (
127	16	53.3	119	2	PSNOA1	phospholipase A2 h
128	16	53.3	119	2	G97003	hypothetical prote
129	16	53.3	120	2	E70424	pillin - Aquifex ae
130	16	53.3	125	2	JH0184	hydrophobin Sc3 pr
131	16	53.3	126	2	D72784	hypothetical prote
132	16	53.3	132	2	I39193	gene HOXA1 protein
133	16	53.3	133	2	B30242	stem cell protein
134	16	53.3	135	2	D90847	hypothetical prote
135	16	53.3	136	2	I50554	gonadotropin II be
136	16	53.3	138	2	E71883	hypothetical prote
137	16	53.3	138	2	A90258	conserved hypothet
138	16	53.3	140	2	T49481	hypothetical prote
139	16	53.3	140	2	T33359	hypothetical prote
140	16	53.3	141	2	T49380	hypothetical prote
141	16	53.3	141	2	PCI1294	trophozoite surfac
142	16	53.3	142	2	J60510	fusaric acid resis
143	16	53.3	143	1	QOEC32	hypothetical prote
144	16	53.3	144	1	PSHUYF	phospholipase A2 (
145	16	53.3	144	2	JN0480	phospholipase A2 (
146	16	53.3	144	2	S01391	phospholipase A2 (
147	16	53.3	144	2	A60480	interleukin-9 prec
148	16	53.3	145	1	PSKRF2U	phospholipase A2 (
149	16	53.3	145	1	S07983	phospholipase A2 h
150	16	53.3	145	2	S01390	phospholipase A2 (
151	16	53.3	145	2	S74292	phospholipase A2 (
152	16	53.3	145	2	T42280	hypothetical prote
153	16	53.3	145	2	T15608	hypothetical prote
154	16	53.3	148	1	PSHU	phospholipase A2 (
155	16	53.3	149	2	C81959	probable pillin NMA
156	16	53.3	150	2	C42744	exonuclease (EC 3.
157	16	53.3	151	2	T25047	hypothetical prote
158	16	53.3	153	2	S34255	hypothetical prote
159	16	53.3	157	2	S55956	probable membrane
160	16	53.3	160	2	JQ0542	185k secretory pro
161	16	53.3	166	2	F84449	hypothetical prote
162	16	53.3	172	2	AB3514	outer membrane pro
163	16	53.3	173	2	T14827	lipid transfer pro
164	16	53.3	174	2	AC0306	probable outer mem
165	16	53.3	177	2	G86917	conserved hypothet
166	16	53.3	177	2	D70654	hypothetical prote
167	16	53.3	178	2	G71195	hypothetical prote
168	16	53.3	181	2	JN0342	N-methyl-D-asparta
169	16	53.3	181	2	T15412	hypothetical prote
170	16	53.3	183	2	JX0311	kunitz type subtitl
171	16	53.3	183	2	I38923	microfibril-associ
172	16	53.3	183	2	A49313	microfibril-associ
173	16	53.3	183	2	S55599	hypothetical prote
174	16	53.3	183	2	S55599	hypothetical prote
175	16	53.3	186	2	B82625	outer membrane pro
176	16	53.3	187	2	S16918	zygote-specific pr
177	16	53.3	187	2	P90315	hypothetical prote
178	16	53.3	188	2	T33886	hypothetical prote
179	16	53.3	189	2	G70347	phenyleacrylic acid
180	16	53.3	190	2	JX0310	kunitz type subtitl
181	16	53.3	190	2	A59416	kunitz type subtitl
182	16	53.3	190	2	A59418	kunitz type subtitl
183	16	53.3	195	2	B82857	plix protein XF003
184	16	53.3	198	2	F83435	conserved hypothet
185	16	53.3	199	2	S35441	serine-cRNA ligase
186	16	53.3	201	2	S23324	gene achaeae prote
187	16	53.3	202	2	AC2875	hypothetical prote
188	16	53.3	204	2	T28939	hypothetical prote
189	16	53.3	205	2	T37947	hypothetical prote
190	16	53.3	207	2	B81418	hypothetical prote
191	16	53.3	208	2	S28675	hypothetical prote
192	16	53.3	208	2	P00278	hypothetical prote
193	16	53.3	208	2	B64925	probable oxidoredu
194	16	53.3	208	2	E90926	probable oxidoredu
195	16	53.3	208	2	A85775	probable oxidoredu
196	16	53.3	213	2	F75585	urease accessory p
197	16	53.3	220	1	S71282	ribosomal protein
198	16	53.3	220	1	T21730	hypothetical prote
199	16	53.3	222	2	I39192	gene HOXA1 protein
200	16	53.3	223	2	A65172	hypothetical prote
201	16	53.3	224	2	A36675	myogenin - mouse
202	16	53.3	225	2	B65127	type 4 prepillin-11
203	16	53.3	225	2	T49769	hypothetical prote
204	16	53.3	226	2	T33361	hypothetical prote
205	16	53.3	227	2	B35882	myogenin - chicken
206	16	53.3	229	1	TRBOTR	trypsin (EC 3.4.21
207	16	53.3	229	2	T29203	hypothetical prote
208	16	53.3	231	1	TRBCTR	trypsin (EC 3.4.21
209	16	53.3	231	2	T02765	glutathione transf
210	16	53.3	233	2	S71464	ribonuclease (EC 3
211	16	53.3	234	2	T52018	fumarate reductase
212	16	53.3	237	2	S08073	cyclic nucleotide
213	16	53.3	237	2	T25152	hypothetical prote
214	16	53.3	238	1	WZBR2	gene 2 protein - h
215	16	53.3	238	2	S31779	trypsin (EC 3.4.21
216	16	53.3	240	2	S39047	trypsin (EC 3.4.21
217	16	53.3	240	2	H82289	conserved hypothet
218	16	53.3	240	2	T25641	hypothetical prote
219	16	53.3	240	2	A43912	myogenin - Japanese
220	16	53.3	241	2	S39048	trypsin (EC 3.4.21
221	16	53.3	242	2	C70895	hypothetical prote
222	16	53.3	242	2	T45795	hypothetical prote
223	16	53.3	245	2	T24565	hypothetical prote
224	16	53.3	247	2	S13813	trypsin (EC 3.4.21
225	16	53.3	247	2	S05494	trypsin (EC 3.4.21
226	16	53.3	248	2	S55067	trypsin (EC 3.4.21
227	16	53.3	248	2	T21786	hypothetical prote
228	16	53.3	248	2	F85067	hypothetical prote
229	16	53.3	249	2	T32060	hypothetical prote
230	16	53.3	254	2	G71545	probable deoxyoctu
231	16	53.3	254	2	A56447	CMR-2-keto-3-deoxy
232	16	53.3	259	2	T32542	hypothetical prote
233	16	53.3	259	2	T46881	hypothetical prote
234	16	53.3	261	2	F82728	succinate dehydrog
235	16	53.3	262	2	F72003	probable outer mem
236	16	53.3	262	2	G86619	probable Omp (impo
237	16	53.3	262	2	F87253	diene lactone hydro
238	16	53.3	262	2	A57391	TcSt-2 protein pre
239	16	53.3	265	2	A95401	protein (imported
240	16	53.3	266	2	S06462	beta-lactamase (EC
241	16	53.3	266	2	S33163	beta-lactamase (EC
242	16	53.3	267	2	S22684	inhibitor of apopt
243	16	53.3	268	2	T10304	apoptosis-inhibiti
244	16	53.3	268	2	A53989	insulin-like growt
245	16	53.3	272	1	A36082	hypothetical prote
246	16	53.3	274	2	T47933	chitinase (EC 3.2.
247	16	53.3	276	2	T09131	diaminopimelate ep
248	16	53.3	279	2	C70458	



249	16	53.3	280	2	T09504	LIM-protein FHL3,	322	16	53.3	366	2	A95382	probable iron upta
250	16	53.3	283	2	A64174	hypothetical prote	323	16	53.3	367	2	T49048	hypothetical prote
251	16	53.3	287	2	T15779	hypothetical prote	324	16	53.3	369	2	AE2345	phospho-N-acetylmu
252	16	53.3	288	2	T04401	endonuclease (EC 3	325	16	53.3	369	2	G81101	membrane-bound lyc
253	16	53.3	288	2	T21790	hypothetical prote	326	16	53.3	369	2	T15213	hypothetical prote
254	16	53.3	288	2	T38075	conserved hypotet	327	16	53.3	370	2	B64632	hydrogenase expres
255	16	53.3	289	1	Q2AD2	early EIA 32K prot	328	16	53.3	370	2	F71883	hydrogenase expres
256	16	53.3	289	1	Q2AD5	hypothetical prote	329	16	53.3	371	2	A39234	opsin - blueottle
257	16	53.3	292	2	T23966	hypothetical prote	330	16	53.3	372	2	C97791	nife protein homol
258	16	53.3	293	2	T03883	hypothetical prote	331	16	53.3	373	1	OORP	rhodopsin - fruit
259	16	53.3	294	2	T23682	hypothetical prote	332	16	53.3	374	2	D83328	probable amniotran
260	16	53.3	296	2	S77427	2-hydroxy-6-oxohep	333	16	53.3	374	2	S40691	opsin rh1 - fruit
261	16	53.3	297	2	A10827	probable phosphog	334	16	53.3	374	2	T05660	hypothetical prote
262	16	53.3	297	2	T38161	hypothetical prote	335	16	53.3	376	2	T31842	zinc finger protei
263	16	53.3	298	2	AF0578	Citg protein (impo	336	16	53.3	376	2	B71652	iron-sulfur cofact
264	16	53.3	302	2	B71310	conserved hypotet	337	16	53.3	377	2	C89456	protein F55A4.8 [i
265	16	53.3	303	2	T19289	hypothetical prote	338	16	53.3	378	2	A82406	cytochrome d ubiq
266	16	53.3	304	2	B83052	rRNA pseudouridine	339	16	53.3	379	2	S67856	Gumg protein - Xan
267	16	53.3	305	2	T49720	hypothetical prote	340	16	53.3	379	2	F71533	conserved hypotet
268	16	53.3	307	1	GSFP3	salivary glue prot	341	16	53.3	384	2	A81678	hypothetical prote
269	16	53.3	309	2	S41427	cysteine proteinas	342	16	53.3	381	2	A35300	G protein-coupled
270	16	53.3	309	2	S37028	exom protein - Rhi	343	16	53.3	383	2	I53870	Edg-1 orphan recep
271	16	53.3	309	2	S39957	succinoglycan bios	344	16	53.3	384	2	S68410	lipase Arab-1 - Ar
272	16	53.3	309	2	C95977	glucosyltransferas	345	16	53.3	384	2	F97711	hypothetical prote
273	16	53.3	311	2	S20085	transcription fact	346	16	53.3	385	2	T49084	hypothetical prote
274	16	53.3	312	1	QOBE07	US10 protein - hum	347	16	53.3	385	2	T04725	hypothetical prote
275	16	53.3	313	2	D85855	probable kinase ye	348	16	53.3	386	2	T06484	aspartate carboxamoy
276	16	53.3	313	2	B91011	probable kinase [i	349	16	53.3	387	2	C41880	nitrogenase cofact
277	16	53.3	313	2	B64985	hypothetical 33.6K	350	16	53.3	388	2	T16861	hypothetical prote
278	16	53.3	314	2	T27686	hypothetical prote	351	16	53.3	389	2	T23167	hypothetical prote
279	16	53.3	314	2	B43652	ferredoxin [2Fe-2S	352	16	53.3	391	1	U01626	attachment protein
280	16	53.3	318	2	B91177	probable hemin per	353	16	53.3	391	2	T08338	conserved hypotet
281	16	53.3	319	2	JC2252	chitinase (EC 3.2.	354	16	53.3	395	2	B47071	nitrogenase cofact
282	16	53.3	321	2	T02235	NBS-IIR type resis	355	16	53.3	399	2	F82657	beta-ketoacyl - (ACP
283	16	53.3	322	2	S54806	chitinase (EC 3.2.	356	16	53.3	399	2	T21015	hypothetical prote
284	16	53.3	323	1	A48149	carcinoma-associat	357	16	53.3	404	2	A46274	HIV gp120-binding
285	16	53.3	323	2	T03614	chitinase (EC 3.2.	358	16	53.3	411	2	E97338	3-oxoacyl - (acyl - ca
286	16	53.3	323	2	S56777	probable	359	16	53.3	412	2	S58105	Cu metalloregulato
287	16	53.3	328	2	C72370	cryptophan-tRNA [i	360	16	53.3	416	2	S27198	homeotic protein H
288	16	53.3	328	2	T36494	probable membrane	361	16	53.3	418	2	T15142	hypothetical prote
289	16	53.3	329	2	JC7935	ankyrin repeat and	362	16	53.3	424	2	UC5891	omega 6 desaturase
290	16	53.3	329	2	AB1327	diaminopimelate ep	363	16	53.3	425	2	B90113	beta1-rRNA synthet
291	16	53.3	329	2	AD1698	diaminopimelate ep	364	16	53.3	425	2	UC5909	AE33 protein - fru
292	16	53.3	330	1	JN0561	urokinase-type pla	365	16	53.3	425	2	A37912	thrombin receptor
293	16	53.3	330	2	I56100	complement factor	366	16	53.3	427	2	G71272	iron-sulfur cofact
294	16	53.3	330	2	C86023	hypothetical prote	367	16	53.3	429	2	T21113	hypothetical prote
295	16	53.3	330	2	T49644	hypothetical prote	368	16	53.3	431	2	T36682	probable secreted
296	16	53.3	331	2	T26004	homeotic protein E	369	16	53.3	432	2	C70381	large subunit of I
297	16	53.3	331	2	A30242	hypothetical prote	370	16	53.3	432	2	A43448	thrombin receptor
298	16	53.3	332	2	AD3535	glucosamine-transpor	371	16	53.3	435	2	T15143	hypothetical prote
299	16	53.3	334	2	H69214	probable thiamin A	372	16	53.3	439	2	T28841	hypothetical prote
300	16	53.3	335	2	B71361	homeobox protein A	373	16	53.3	442	2	T24196	hypothetical prote
301	16	53.3	335	2	G01448	effector protein H	374	16	53.3	442	2	D84600	probable xylose
302	16	53.3	337	2	A53041	effector cell prot	375	16	53.3	444	2	F86329	hypothetical prote
303	16	53.3	338	2	G88986	protein C50H11.5 [	376	16	53.3	448	2	S41725	integrinase - Saccha
304	16	53.3	338	2	T06603	hypothetical prote	377	16	53.3	449	2	T26571	hypothetical prote
305	16	53.3	340	2	T48288	hypothetical prote	378	16	53.3	450	2	T17234	hypothetical prote
306	16	53.3	344	1	RMRTC2	T-cell surface gly	379	16	53.3	450	2	AD3117	polygalacturonase
307	16	53.3	344	1	G82428	iron(III) ABC tran	380	16	53.3	450	2	A96170	hypothetical prote
308	16	53.3	344	2	S72723	dipeptide transpor	381	16	53.3	456	2	T31857	ribonuclease inhib
309	16	53.3	345	2	B85881	hypothetical prote	382	16	53.3	456	2	T37575	bacterial cell cyc
310	16	53.3	345	2	B65012	hypothetical prote	383	16	53.3	466	2	AD3477	malate dehydrogena
311	16	53.3	345	2	H91036	hypothetical prote	384	16	53.3	470	2	A40697	63K sperm flagella
312	16	53.3	345	2	A05279	surface antigen 51	385	16	53.3	474	2	S31712	beta-1,3-glucanase
313	16	53.3	347	2	S41638	T-cell surface gly	386	16	53.3	475	2	S53389	glutamate decarbox
314	16	53.3	348	2	JU0431	hypothetical 35.5K	387	16	53.3	475	2	T50686	peroxisomal Ca-dep
315	16	53.3	349	2	S53864	hypothetical prote	388	16	53.3	477	1	I38893	transcription fact
316	16	53.3	349	2	T16882	hypothetical prote	389	16	53.3	478	2	S31906	beta-1,3-glucanase
317	16	53.3	355	2	C96651	protein Ttp18.9 [i	390	16	53.3	481	2	T20309	stretch-activated
318	16	53.3	355	2	T00578	probable GDSL-moti	391	16	53.3	486	2	JC7212	hypothetical prote
319	16	53.3	359	2	D89007	protein F59B1.6 [i	392	16	53.3	486	2	T38087	hypothetical prote
320	16	53.3	359	2	T26813	hypothetical prote	393	16	53.3	487	2	F84727	hypothetical prote
321	16	53.3	364	2	B36313	hypothetical 42K p	394	16	53.3	488	1	JC2385	protein disulfide-

395	16	53.3	492	2	JC5169	alkaline nuclease
396	16	53.3	496	2	D83614	conserved hypotet
397	16	53.3	498	2	AF0157	prolyl 4-hydroxylase
398	16	53.3	498	2	A48203	interleukin-14 pre
399	16	53.3	499	2	B81914	probable periplasm
400	16	53.3	504	1	A28807	protein disulfide-
401	16	53.3	504	2	S41661	protein disulfide-
402	16	53.3	504	2	G02474	interferon regulat
403	16	53.3	505	1	B69417	proteobacter thymidine
404	16	53.3	505	1	S55507	protein disulfide-
405	16	53.3	505	2	S63994	protein disulfide-
406	16	53.3	505	2	S68363	protein disulfide-
407	16	53.3	505	2	JC5704	protein disulfide-
408	16	53.3	509	2	S46314	hydroxymethylgluta
409	16	53.3	509	2	H70597	probable membrane
410	16	53.3	510	2	T48977	carboxypeptidase-1
411	16	53.3	510	2	H90787	Rtn-like protein [
412	16	53.3	512	2	T37819	probable zinc meta
413	16	53.3	514	2	G01026	serine-tRNA ligase
414	16	53.3	514	2	D56849	dopamine receptor-
415	16	53.3	519	2	T45764	hypothetical prote
416	16	53.3	521	2	A99100	phosphatidylinosit
417	16	53.3	521	2	T27192	hypothetical prote
418	16	53.3	522	2	S71821	probable interleuk
419	16	53.3	522	2	T13687	hypothetical prote
420	16	53.3	528	2	T33599	hypothetical prote
421	16	53.3	533	2	H71492	probable hep-60 -
422	16	53.3	534	2	B81654	60 kDa chaperonin
423	16	53.3	536	2	H85647	hypothetical prote
424	16	53.3	547	1	A32244	60K cysteine-rich
425	16	53.3	547	2	S19607	alkaline phosphata
426	16	53.3	547	2	B43584	60K cysteine-rich
427	16	53.3	548	2	B84306	hypothetical prote
428	16	53.3	553	2	D71515	60K cysteine-rich
429	16	53.3	554	2	C81671	60 kDa outer membr
430	16	53.3	554	2	T25288	hypothetical prote
431	16	53.3	556	1	S12602	60K cysteine-rich
432	16	53.3	556	2	A86560	60 kDa Cysteine-ri
433	16	53.3	557	1	B39439	60K cysteine-rich
434	16	53.3	557	2	D84496	probable retroelem
435	16	53.3	557	2	A48434	variant-specific s
436	16	53.3	558	2	JC5204	60K cysteine-rich
437	16	53.3	559	2	S63388	probable membrane
438	16	53.3	569	2	D85354	hypothetical prote
439	16	53.3	573	2	T49610	related to rna bin
440	16	53.3	577	2	T28017	hypothetical prote
441	16	53.3	579	2	T24880	hypothetical prote
442	16	53.3	589	2	T43210	fibulin-1D precurs
443	16	53.3	591	2	T39195	probable amino aci
444	16	53.3	600	2	B46642	DNA-directed DNA p
445	16	53.3	604	2	B96744	unknown protein [I
446	16	53.3	604	2	T49577	hypothetical prote
447	16	53.3	617	2	T49444	lustrin A related
448	16	53.3	631	2	S71508	dnak-type molecula
449	16	53.3	638	2	G02068	white homolog - hu
450	16	53.3	643	2	T25473	hypothetical prote
451	16	53.3	655	2	T51792	hypothetical prote
452	16	53.3	656	2	JC2005	integrin beta-5 ch
453	16	53.3	657	2	E72486	hypothetical prote
454	16	53.3	668	2	A42375	chemoreceptor prot
455	16	53.3	672	2	T12524	hypothetical prote
456	16	53.3	677	2	C42125	trophozoite cystel
457	16	53.3	680	2	PN0510	integrin beta-3 ch
458	16	53.3	686	2	S30075	ferric reductase (
459	16	53.3	686	2	S43562	KO8B5.3 protein -
460	16	53.3	689	2	T42760	fibulin, splice fo
461	16	53.3	692	2	T32980	hypothetical prote
462	16	53.3	704	2	S21911	BRcore-NS-73 prote
463	16	53.3	704	2	T31227	cran protein homol
464	16	53.3	706	2	A81848	probable TonB-depe
465	16	53.3	707	2	G86894	hypothetical prote
466	16	53.3	709	2	T29692	hypothetical prote
467	16	53.3	712	2	T42990	fibulin 1, splice
468	16	53.3	713	2	I50128	fibroblast growth
469	16	53.3	713	2	A35502	major surface-labe
470	16	53.3	722	2	E71403	hypothetical prote
471	16	53.3	723	2	PN0509	integrin beta-3 ch
472	16	53.3	734	2	JC4861	ferriillin beta cha
473	16	53.3	735	2	G02937	ferriillin beta - cr
474	16	53.3	743	2	T02147	hypothetical prote
475	16	53.3	746	2	S62365	SNR1-related prote
476	16	53.3	752	2	T48574	hypothetical prote
477	16	53.3	752	2	T20871	hypothetical prote
478	16	53.3	753	2	B36268	platelet glycoprot
479	16	53.3	758	2	T48815	mixed-linked glucan
480	16	53.3	761	2	A46193	88K E-26-specific
481	16	53.3	763	2	E96693	probable terpene s
482	16	53.3	767	2	T07693	hypothetical prote
483	16	53.3	769	2	S50966	probable membrane
484	16	53.3	770	2	I40027	nitrite reductase
485	16	53.3	773	2	I46059	beta-1 integrin su
486	16	53.3	778	2	A60798	platelet glycoprot
487	16	53.3	781	2	S43534	integrin beta3 - C
488	16	53.3	788	2	A26547	platelet glycoprot
489	16	53.3	788	2	I77349	platelet glycoprot
490	16	53.3	788	2	I51530	integrin beta-3 su
491	16	53.3	798	2	B27079	fibronectin recept
492	16	53.3	798	2	A28193	integrin beta-1 ch
493	16	53.3	798	2	B28193	integrin beta-1* C
494	16	53.3	799	2	T22793	hypothetical prote
495	16	53.3	799	2	A36308	integrin beta-5 ch
496	16	53.3	801	2	G82302	probable cellobios
497	16	53.3	805	2	I40029	nitrite reductase
498	16	53.3	807	2	E83726	assimilatory nitri
499	16	53.3	810	2	B30848	plasma (EC 3.4.21
500	16	53.3	817	2	T25674	hypothetical prote
501	16	53.3	818	2	JC4058	fibroblast growth
502	16	53.3	822	2	S19947	fibroblast growth
503	16	53.3	822	2	B49151	fibroblast growth
504	16	53.3	822	2	T01622	probable salt-indu
505	16	53.3	823	2	D66165	protein F15K9.3 (f
506	16	53.3	829	2	JC4583	fibroblast growth
507	16	53.3	834	2	T09369	hypothetical prote
508	16	53.3	838	2	T04449	hypothetical prote
509	16	53.3	849	2	E86306	Similar to tuftell
510	16	53.3	849	2	T01286	probable RNA-bind
511	16	53.3	863	1	S51789	VLDR receptor prec
512	16	53.3	872	2	S33015	hypothetical prote
513	16	53.3	884	2	E75489	conserved hypotet
514	16	53.3	885	2	JN0339	N-methyl-D-asparta
515	16	53.3	889	2	D71414	hypothetical prote
516	16	53.3	897	2	S67283	hypothetical prote
517	16	53.3	901	2	JN0337	N-methyl-D-asparta
518	16	53.3	905	2	T23229	hypothetical prote
519	16	53.3	906	2	JN0341	N-methyl-D-asparta
520	16	53.3	907	2	A46296	N-methyl-D-asparta
521	16	53.3	907	2	E96692	probable wall-asso
522	16	53.3	913	2	G64110	hypothetical prote
523	16	53.3	922	2	JN0340	N-methyl-D-asparta
524	16	53.3	922	2	JN0338	N-methyl-D-asparta
525	16	53.3	936	2	S64384	probable membrane
526	16	53.3	938	2	A46612	N-methyl-D-asparta
527	16	53.3	938	2	S21104	N-methyl-D-asparta
528	16	53.3	938	2	S19710	N-methyl-D-asparta
529	16	53.3	943	2	A47551	N-methyl-D-asparta
530	16	53.3	946	2	T01460	hypothetical prote
531	16	53.3	959	2	JN0336	N-methyl-D-asparta
532	16	53.3	964	2	S48404	probable membrane
533	16	53.3	991	1	PIBVCV	RNA I protein - cu
534	16	53.3	992	2	A31666	hypothetical prote
535	16	53.3	1007	2	T01437	hypothetical prote
536	16	53.3	1025	2	B54718	dihydropyrimidine
537	16	53.3	1025	2	A54718	dihydropyrimidine
538	16	53.3	1029	2	S56229	probable membrane
539	16	53.3	1044	2	T43155	nitrite reductase
540	16	53.3	1044	2	T43160	nitrite reductase

541	16	53.3	1046	2	A26838	614	15	50.0	27	2	PC4234	hypothetical prote
542	16	53.3	1069	2	S27922	615	15	50.0	31	2	B61014	defensin R-2 - rat
543	16	53.3	1084	2	T15616	616	15	50.0	34	2	G44336	neurotoxin Tx3-3 -
544	16	53.3	1099	2	T16283	617	15	50.0	35	2	I18224	filamentous hemaggl
545	16	53.3	1101	2	T16840	618	15	50.0	40	2	UT0515	Ig heavy chain V-I
546	16	53.3	1129	2	A47511	619	15	50.0	42	2	A37907	thyrotropin beta c
547	16	53.3	1132	2	JW0091	620	15	50.0	45	2	AC3539	hypothetical prote
548	16	53.3	1132	2	JC4127	621	15	50.0	46	2	B81072	hypothetical prote
549	16	53.3	1153	2	P84468	622	15	50.0	46	2	G53613	plectoxin IX - spi
550	16	53.3	1218	2	P84537	623	15	50.0	52	1	XBP1	bromelain inhibito
551	16	53.3	1235	2	C69165	624	15	50.0	52	2	S66609	bromelain inhibito
552	16	53.3	1236	2	T50904	625	15	50.0	53	2	T10299	conotoxin-like pro
553	16	53.3	1274	2	T50729	626	15	50.0	53	2	T10405	conotoxin-like pro
554	16	53.3	1297	2	T52065	627	15	50.0	53	2	C72850	conotoxin homolog
555	16	53.3	1297	2	T30274	628	15	50.0	53	2	T30499	conotoxin-like pro
556	16	53.3	1307	2	G96711	629	15	50.0	54	1	FEME	ferredoxin 2[Fe-4
557	16	53.3	1316	2	G70535	630	15	50.0	56	1	TIPIG	pancreatic secreto
558	16	53.3	1321	2	JE0352	631	15	50.0	57	1	NILLSA	nisin precursor -
559	16	53.3	1336	2	T23310	632	15	50.0	58	2	D82759	hypothetical prote
560	16	53.3	1347	2	T41321	633	15	50.0	59	2	T44147	B3 protein [import
561	16	53.3	1357	2	T16860	634	15	50.0	59	2	AE3319	hypothetical prote
562	16	53.3	1360	2	T33922	635	15	50.0	62	2	P82332	bacterioferritin-a
563	16	53.3	1373	2	JE0095	636	15	50.0	63	2	S25772	testis-specific pr
564	16	53.3	1376	2	S63986	637	15	50.0	63	2	A34484	metallothionein 1
565	16	53.3	1396	1	VCBB40	638	15	50.0	64	2	AD2881	hypothetical prote
566	16	53.3	1397	2	T46354	639	15	50.0	66	2	T17934	hypothetical prote
567	16	53.3	1413	2	D84481	640	15	50.0	68	2	B84730	hypothetical prote
568	16	53.3	1416	2	P88550	641	15	50.0	69	2	B97230	hypothetical prote
569	16	53.3	1444	2	T18856	642	15	50.0	70	2	S23316	hypothetical prote
570	16	53.3	1469	2	B36665	643	15	50.0	70	2	C83620	hypothetical prote
571	16	53.3	1475	2	A60026	644	15	50.0	71	2	F69981	hypothetical prote
572	16	53.3	1509	2	T19486	645	15	50.0	71	2	T00008	copy number contro
573	16	53.3	1518	2	T28880	646	15	50.0	71	2	A59412	KGD-bearing plate
574	16	53.3	1523	2	T13953	647	15	50.0	71	2	A59413	platelet-aggregati
575	16	53.3	1531	2	T42218	648	15	50.0	71	2	S13168	batroxostatin - ba
576	16	53.3	1557	2	T28811	649	15	50.0	72	2	G43019	platelet aggregati
577	16	53.3	1559	2	T16210	650	15	50.0	72	2	A42325	ori 5' to pheC - p
578	16	53.3	1620	2	T27283	651	15	50.0	72	2	AB3528	hypothetical prote
579	16	53.3	1650	2	S53457	652	15	50.0	72	2	F43019	platelet aggregati
580	16	53.3	1700	2	S08167	653	15	50.0	72	2	D43019	platelet aggregati
581	16	53.3	1772	2	A45532	654	15	50.0	72	2	T43019	platelet aggregati
582	16	53.3	1784	2	T02844	655	15	50.0	72	2	B43019	platelet aggregati
583	16	53.3	1801	1	MMRPS	656	15	50.0	72	2	B43020	platelet aggregati
584	16	53.3	1822	2	S63985	657	15	50.0	73	2	T42302	hypothetical prote
585	16	53.3	1896	2	T01490	658	15	50.0	73	2	S78719	protein YER091c-a
586	16	53.3	2022	2	T48818	659	15	50.0	73	2	H43019	platelet aggregati
587	16	53.3	2049	2	T29227	660	15	50.0	73	2	C43019	platelet aggregati
588	16	53.3	2180	2	T29764	661	15	50.0	73	2	B43019	platelet aggregati
589	16	53.3	2182	2	T14320	662	15	50.0	73	2	A43020	platelet aggregati
590	16	53.3	2233	2	T28669	663	15	50.0	73	2	A43019	platelet aggregati
591	16	53.3	2287	2	T23312	664	15	50.0	73	2	B40003	platelet aggregati
592	16	53.3	2399	2	H71879	665	15	50.0	73	2	A40003	platelet aggregati
593	16	53.3	2470	2	I50726	666	15	50.0	74	2	S33833	hypothetical prote
594	16	53.3	2529	2	B64635	667	15	50.0	74	2	H90820	hypothetical prote
595	16	53.3	2531	2	T16743	668	15	50.0	75	2	S19696	mating pheromone B
596	16	53.3	2548	2	E59435	669	15	50.0	76	1	TIAM3	protease inhibitor
597	16	53.3	2643	2	T29149	670	15	50.0	76	2	I83048	FSH beta-subunit -
598	16	53.3	2704	2	S09118	671	15	50.0	76	2	G68226	hypothetical prote
599	16	53.3	2718	2	A23475	672	15	50.0	76	2	A33034	neurogranin - bovi
600	16	53.3	2761	2	T21064	673	15	50.0	76	2	D44007	apoptoxin IV - tra
601	16	53.3	3133	2	S52093	674	15	50.0	76	2	C44007	apoptoxin VI - tra
602	16	53.3	3672	2	T23433	675	15	50.0	76	2	AH2120	hypothetical prote
603	16	53.3	3704	2	T37316	676	15	50.0	77	2	JN0536	hypothetical 8.3K
604	16	53.3	4056	2	H96599	677	15	50.0	78	2	D84620	hypothetical prote
605	16	53.3	4344	1	S02392	678	15	50.0	78	2	G98261	hypothetical prote
606	16	53.3	4545	1	S25111	679	15	50.0	79	2	C60767	retrovirus-related
607	16	53.3	4548	1	S00657	680	15	50.0	79	2	AH3369	hypothetical prote
608	16	53.3	4660	2	T42737	681	15	50.0	80	2	I65235	testicular lutealni
609	16	53.3	4753	1	A47437	682	15	50.0	81	2	C65019	hypothetical prote
610	16	53.3	5107	2	T29144	683	15	50.0	81	2	AC1864	hypothetical prote
611	16	53.3	10797	2	T30192	684	15	50.0	83	1	WBBP67	gene 18.7 protein
612	16	50.0	15	2	B59045	685	15	50.0	83	2	F35982	bitan alpha - puff
613	15	50.0	27	2	S55030	686	15	50.0	83	2	A34156	bitistatin - puff

687	15	50.0	85	1	W9BPB7	gene 19.2 protein	760	15	50.0	119	2	PH1554	Ig H chain V regio
688	15	50.0	85	2	A12765	hypothetical prote	761	15	50.0	119	2	A45574	13.3 kDa polyepi
689	15	50.0	86	2	S20471	class V zygote-spe	762	15	50.0	120	1	PSKFT3U	phospholipase A2 (
690	15	50.0	87	2	T18115	hypothetical prote	763	15	50.0	120	1	PSKFT3U	phospholipase A2 (
691	15	50.0	88	2	F89129	protein F52EL.6 (l	764	15	50.0	120	2	I50993	gonadotropin i bet
692	15	50.0	89	2	F90453	ferredoxin like pr	765	15	50.0	120	2	G71105	hypothetical prote
693	15	50.0	91	2	C38180	hypothetical prote	766	15	50.0	120	2	A12181	hypothetical prote
694	15	50.0	92	2	B90856	anther-specific pr	767	15	50.0	121	1	PSEBAC	phospholipase A2 (
695	15	50.0	94	2	T03285	prostatic steroid	768	15	50.0	121	1	RMBEM3	surfact glycoprote
696	15	50.0	95	1	B0R3	3-mercaptopyruvate	769	15	50.0	121	1	PH1661	Ig heavy chain V r
697	15	50.0	96	2	A57483	alpha-(l)-3-fucosy	770	15	50.0	122	1	PSABA	phospholipase A2 (
698	15	50.0	97	2	B97832	hypothetical prote	771	15	50.0	122	1	PSR5AE	phospholipase A2 (
699	15	50.0	98	2	T33302	hypothetical prote	772	15	50.0	122	1	PSR5AW	phospholipase A2 (
700	15	50.0	98	2	S42596	hypothetical prote	773	15	50.0	122	1	PSYIT1	phospholipase A2 i
701	15	50.0	100	2	G87427	hypothetical prote	774	15	50.0	122	2	A53872	phospholipase A2 (
702	15	50.0	101	1	MMV211	11k protein - vacc	775	15	50.0	122	2	A25806	phospholipase A2 (
703	15	50.0	101	1	MMV212	11k protein - vacc	776	15	50.0	122	2	A44179	phospholipase A2 (
704	15	50.0	101	2	T30793	probable DNA-bind	777	15	50.0	122	2	S46979	phospholipase A2 -
705	15	50.0	101	2	S73700	probable lipoprote	778	15	50.0	122	2	S62780	phospholipase A2 (
706	15	50.0	102	2	S50530	hypothetical prote	779	15	50.0	122	2	UX0063	phospholipase A2 (
707	15	50.0	102	2	S69884	hypothetical prote	780	15	50.0	122	2	S09314	phospholipase A2 (
708	15	50.0	102	2	C81889	hypothetical prote	781	15	50.0	122	2	T49039	response regulator
709	15	50.0	102	2	B82751	hypothetical prote	782	15	50.0	123	2	B53872	phospholipase A2 (
710	15	50.0	103	2	C69000	hypothetical prote	783	15	50.0	124	1	PSBGA2	phospholipase A2 (
711	15	50.0	103	2	T58936	hypothetical prote	784	15	50.0	124	1	NRCM	pancreatic ribonuc
712	15	50.0	104	2	D71035	hypothetical prote	785	15	50.0	124	1	NRCMB	pancreatic ribonuc
713	15	50.0	104	2	E81283	hypothetical prote	786	15	50.0	124	1	NRCM	pancreatic ribonuc
714	15	50.0	104	2	S50911	metallochionein-2	787	15	50.0	124	1	A26535	phospholipase A2 (
715	15	50.0	106	2	S74084	foliitropin beta c	788	15	50.0	124	2	A59420	phospholipase A2 (
716	15	50.0	106	2	E69048	conserved hypotet	789	15	50.0	124	2	I51190	phospholipase A2 -
717	15	50.0	106	2	S32032	Sp17 protein precu	790	15	50.0	124	2	S25113	insulin-like growt
718	15	50.0	108	2	T45063	hypothetical prote	791	15	50.0	125	2	UX0052	phospholipase A2 (
719	15	50.0	109	2	PH1653	Ig heavy chain V r	792	15	50.0	125	2	S65624	phospholipase A(2)
720	15	50.0	109	2	T46171	hypothetical prote	793	15	50.0	125	2	AD0297	probable membrane
721	15	50.0	109	2	E70956	hypothetical prote	794	15	50.0	125	2	T28743	hypothetical prote
722	15	50.0	110	2	S55208	hypothetical prote	795	15	50.0	126	2	I36914	CD59 protein - bab
723	15	50.0	111	1	A29654	proteinase inhibit	796	15	50.0	126	2	F84240	hypothetical prote
724	15	50.0	111	1	S69911	Ig V-D-J region (R	797	15	50.0	126	2	G90350	conserved hypotet
725	15	50.0	111	2	H91187	hypothetical prote	798	15	50.0	127	2	S40034	Killer toxin KP4 p
726	15	50.0	111	2	H85046	hypothetical prote	799	15	50.0	127	2	T10546	Rab11 protein homo
727	15	50.0	112	1	B29654	proteinase inhibit	800	15	50.0	128	1	RWHU59	surface glycoprote
728	15	50.0	113	1	S07092	gonadotropin beta	801	15	50.0	128	2	I36894	CD59 protein - gre
729	15	50.0	113	2	T36789	hypothetical prote	802	15	50.0	128	2	T35073	probable phosphori
730	15	50.0	115	2	S66924	probable membrane	803	15	50.0	128	2	AH3422	hypothetical membr
731	15	50.0	115	2	T45252	phosphoribosyl-AMP	804	15	50.0	129	1	FTTUB	foliitropin beta c
732	15	50.0	117	2	A34860	phospholipase A2 (	805	15	50.0	129	1	A23550	foliitropin beta c
733	15	50.0	117	2	B32416	phospholipase A2 (	806	15	50.0	129	2	S73514	foliitropin beta c
734	15	50.0	117	2	A32416	phospholipase A2 (	807	15	50.0	129	2	JC4526	probable lipoprote
735	15	50.0	117	2	T26924	hypothetical prote	808	15	50.0	130	2	A33893	foliitropin beta c
736	15	50.0	118	2	PSKFT1	phospholipase A2 (	809	15	50.0	130	2	A33893	foliitropin beta c
737	15	50.0	118	1	PSKFT2	phospholipase A2 (	810	15	50.0	131	2	A46298	complement regulat
738	15	50.0	118	1	PSKFT3	phospholipase A2 (	811	15	50.0	131	2	A46298	pigment depositio
739	15	50.0	118	1	PSOXB	phospholipase A2 h	812	15	50.0	131	2	A83350	hypothetical prote
740	15	50.0	118	1	FTHOB	foliitropin beta c	813	15	50.0	132	1	PSHOA	phospholipase A2 (
741	15	50.0	118	2	A32622	phospholipase A2 (	814	15	50.0	132	2	I37143	agouti protein pre
742	15	50.0	118	2	B34860	phospholipase A2 (	815	15	50.0	132	2	B83382	hypothetical prote
743	15	50.0	118	2	C34860	phospholipase A2 (	816	15	50.0	133	1	PSOXG	phospholipase A2 (
744	15	50.0	118	2	S01801	phospholipase A2 (	817	15	50.0	133	2	T1R2BR	crystin inhibitor
745	15	50.0	118	2	PN0141	lutropin beta chai	818	15	50.0	133	2	S14728	textilotoxin chain
746	15	50.0	118	2	PN0139	lutropin beta chai	819	15	50.0	134	2	H84868	hypothetical prote
747	15	50.0	118	2	S61051	hypothetical prote	820	15	50.0	135	1	C64872	hypothetical prote
748	15	50.0	118	2	T49515	hypothetical prote	821	15	50.0	135	2	E83461	hypothetical prote
749	15	50.0	119	1	T32397	hypothetical prote	822	15	50.0	135	2	T15610	hypothetical prote
750	15	50.0	119	1	PSNJ2K	phospholipase A2 (	823	15	50.0	136	2	AB0846	formate hydrogenu
751	15	50.0	119	1	PSNJ3K	phospholipase A2 (	824	15	50.0	136	2	S08626	formate hydrogenu
752	15	50.0	119	1	PSBYA	phospholipase A2 (	825	15	50.0	136	2	F91075	formate hydrogenu
753	15	50.0	119	1	PSNJ2W	phospholipase A2 (	826	15	50.0	136	2	H82787	hypothetical prote
754	15	50.0	119	1	PSNJ3W	phospholipase A2 (	827	15	50.0	136	2	T23996	hypothetical prote
755	15	50.0	119	1	PSOXA	phospholipase A2 (	828	15	50.0	136	2	T36241	hypothetical prote
756	15	50.0	119	2	JN0403	phospholipase A2 h	829	15	50.0	137	1	JC4877	hypothetical prote
757	15	50.0	119	2	S07528	phospholipase A2 (	830	15	50.0	137	2	JC4710	beta-bungarotoxin
758	15	50.0	119	2	JX0115	pancreatic ribonuc	831	15	50.0	137	2	S37848	hypothetical prote
759	15	50.0	119	2	A61465	lutropin beta chai	832	15	50.0	137	2	T49243	hypothetical prote

833	15	50.0	137	2	T15609	hypothetical prote	906	15	50.0	146	2	I48342	phospholipase A2 (
834	15	50.0	137	2	A96914	uncharacterized pr	907	15	50.0	146	2	T01683	immunobiliaction ant
835	15	50.0	137	2	AE2763	transcription regu	908	15	50.0	147	1	PSKPA4	phospholipase A2 (
836	15	50.0	137	2	AF2790	conserved hypothet	909	15	50.0	147	2	I52489	ribonuclease 4 (EC
837	15	50.0	137	2	D97544	transcription regu	910	15	50.0	147	2	A53180	ribonuclease PL3 (
838	15	50.0	138	1	I51380	phospholipase A2 h	911	15	50.0	147	2	I50994	gonadotropin II be
839	15	50.0	138	1	PSRSAT	phospholipase A2 h	912	15	50.0	147	2	T20563	hypothetical prote
840	15	50.0	138	1	PSVIF	phospholipase A2 (	913	15	50.0	147	2	E82523	hypothetical prote
841	15	50.0	138	1	PSVIAA	phospholipase A2 (	914	15	50.0	147	2	F70216	hypothetical prote
842	15	50.0	138	1	PSVIAC	phospholipase A2 (	915	15	50.0	148	2	H82065	PTS system, nitrog
843	15	50.0	138	1	BGHU2	spermatid transiti	916	15	50.0	148	2	I47180	Ig heavy chain var
844	15	50.0	138	2	A48168	phospholipase A2 (	917	15	50.0	149	2	T01676	cytochrome-c oxida
845	15	50.0	138	2	A46169	phospholipase A2 (	918	15	50.0	149	2	S70979	bifl. protein - Esc
846	15	50.0	138	2	F48188	phospholipase A2 (	919	15	50.0	149	2	AG2708	conserved hypothet
847	15	50.0	138	2	S29299	phospholipase A2 (	920	15	50.0	150	2	F69883	conserved hypothet
848	15	50.0	138	2	S59522	phospholipase A2 (	921	15	50.0	151	2	C90048	conserved hypothet
849	15	50.0	138	2	S10333	ammodontoxin B prec	922	15	50.0	152	2	G81161	hypothetical prote
850	15	50.0	138	2	I51386	phospholipase A2 (	923	15	50.0	153	2	T13054	hypothetical prote
851	15	50.0	138	2	S10992	ammodytin L precur	924	15	50.0	153	2	PN0103	hypothetical 17K p
852	15	50.0	138	2	JC1342	phospholipase A2 (	925	15	50.0	154	2	S35207	hypothetical 7 - buf
853	15	50.0	138	2	I50098	phospholipase a2 -	926	15	50.0	154	2	S23320	hypothetical prote
854	15	50.0	138	2	A29479	thytotropin beta c	927	15	50.0	154	2	T27967	hypothetical prote
855	15	50.0	138	2	PC1211	envelope protein -	928	15	50.0	156	2	T31839	hypothetical prote
856	15	50.0	138	2	PC1212	envelope protein -	929	15	50.0	156	2	S74049	hypothetical prote
857	15	50.0	138	2	PC1205	envelope protein -	930	15	50.0	157	2	G65156	putative electron
858	15	50.0	138	2	PC1203	envelope protein -	931	15	50.0	157	2	S58017	hypothetical prote
859	15	50.0	138	2	PC1200	genome polyprotein	932	15	50.0	157	2	S58338	hypothetical prote
860	15	50.0	138	2	PC1210	envelope protein -	933	15	50.0	158	2	T52365	hypothetical prote
861	15	50.0	138	2	PC1207	envelope protein -	934	15	50.0	158	2	G86964	hypothetical prote
862	15	50.0	138	2	PC1197	genome polyprotein	935	15	50.0	158	2	T22829	hypothetical prote
863	15	50.0	138	2	PC1209	envelope protein -	936	15	50.0	159	2	S46822	hypothetical prote
864	15	50.0	138	2	PC1199	genome polyprotein	937	15	50.0	159	2	H81657	conserved hypothet
865	15	50.0	138	2	PC1201	envelope protein -	938	15	50.0	159	2	D71500	hypothetical prote
866	15	50.0	138	2	T14185	chitinase (EC 3.2.	939	15	50.0	159	2	T33966	hypothetical prote
867	15	50.0	138	2	T21792	hypothetical prote	940	15	50.0	160	2	T23995	hypothetical prote
868	15	50.0	139	1	PSIV	phospholipase A2 (	941	15	50.0	161	2	H64504	hypothetical prote
869	15	50.0	139	1	I52320	testicular luteini	942	15	50.0	161	2	S40711	hypothetical prote
870	15	50.0	139	2	A38612	insulin-like growt	943	15	50.0	161	2	T48146	disulfide bond for
871	15	50.0	139	2	S31612	beta-1,3-glucanase	944	15	50.0	162	2	H81058	hypothetical prote
872	15	50.0	139	2	AG0578	conserved hypothet	945	15	50.0	162	2	T50253	hypothetical prote
873	15	50.0	140	1	I1MSG1	interleukin-4 prec	946	15	50.0	162	2	T18188	hypothetical prote
874	15	50.0	140	1	HBFGC	hemoglobin beta ch	947	15	50.0	162	2	A72738	hypothetical prote
875	15	50.0	140	2	A48166	gonadotropin II be	948	15	50.0	162	2	E86387	18 .7K hypothetical
876	15	50.0	140	2	H69405	hypothetical prote	949	15	50.0	163	2	A71469	hypothetical prote
877	15	50.0	140	2	T29914	hypothetical prote	950	15	50.0	164	1	A91016	ferredoxin-type pr
878	15	50.0	140	2	E87035	hypothetical prote	951	15	50.0	164	1	C85860	ferredoxin-type pr
879	15	50.0	141	1	B60826	gonadotropin beta	952	15	50.0	164	1	F64990	ferredoxin-type pr
880	15	50.0	141	1	UTBOB	lutropin beta chai	953	15	50.0	164	2	C35216	tumor suppressor p
881	15	50.0	141	1	UTBOB	lutropin beta chai	954	15	50.0	164	2	I84725	choriogonadotropin
882	15	50.0	141	1	UTRCH	lutropin beta chai	955	15	50.0	165	1	KTBAB	pepidoglycan asno
883	15	50.0	141	1	UTRCH	lutropin beta chai	956	15	50.0	165	1	KTBAB	hypothetical prote
884	15	50.0	141	1	UTSHB	lutropin beta chai	957	15	50.0	165	2	I40769	hypothetical prote
885	15	50.0	141	2	UC4527	luteinizing hormon	958	15	50.0	165	2	H84634	hypothetical prote
886	15	50.0	142	2	S54243	Ig mu heavy chain	959	15	50.0	165	2	E84346	hypothetical prote
887	15	50.0	142	2	AD2696	lipa protein (impo	960	15	50.0	165	2	S59899	chlorocorturin chal
888	15	50.0	142	2	E96711	F24U5 .7 (imported)	961	15	50.0	166	2	JN0248	placlet-derived g
889	15	50.0	143	2	T20951	hypothetical prote	962	15	50.0	166	2	JC1348	hypothetical 18K p
890	15	50.0	144	1	UTCAB	gonadotropin beta	963	15	50.0	166	2	A40783	18 .2K protein - ph
891	15	50.0	144	1	JL0089	interleukin-9 prec	964	15	50.0	170	2	C96009	probable plasmid s
892	15	50.0	144	2	E88485	proteine F23F12 .3 (	965	15	50.0	171	2	A10607	probable lipoprote
893	15	50.0	144	2	T16824	hypothetical prote	966	15	50.0	171	2	D90747	probable enzyme (l
894	15	50.0	145	1	PSBOA	phospholipase A2 (	967	15	50.0	171	2	A64825	Ybp protein precu
895	15	50.0	145	1	PSKRA2	phospholipase A2 (	968	15	50.0	171	2	T33072	hypothetical prote
896	15	50.0	145	2	I48093	phospholipase A2 (	969	15	50.0	171	2	E85046	hypothetical prote
897	15	50.0	145	2	I37231	beta-gonadotropin	970	15	50.0	172	2	T20210	hypothetical prote
898	15	50.0	146	1	PSDG	phospholipase A2 (	971	15	50.0	172	2	T10346	hypothetical prote
899	15	50.0	146	1	PSNJA'	phospholipase A2 (	972	15	50.0	173	1	B45932	atrial gland pepti
900	15	50.0	146	1	PSPCA	phospholipase A2 (	973	15	50.0	173	1	GOGAAA	atrial gland pepti
901	15	50.0	146	1	S34049	phospholipase A2 (	974	15	50.0	173	1	GOGABA	atrial gland pepti
902	15	50.0	146	1	S16763	gonadotropin beta	975	15	50.0	173	2	S62349	L71-3 protein - fr
903	15	50.0	146	2	A35493	phospholipase A2 (	976	15	50.0	173	2	H85597	probable enzyme yb
904	15	50.0	146	2	A33394	phospholipase A2 (	977	15	50.0	173	2	E97569	hypothetical prote
905	15	50.0	146	2	JU0283	phospholipase A2 (	978	15	50.0	174	2	T15176	hypothetical prote

979 15 50.0 174 2 A86358 similar to blue co  
980 15 50.0 175 2 T48541 hypothetical prote  
981 15 50.0 176 2 T18921 hypothetical prote  
982 15 50.0 177 1 AFMD8 phycoerythrin beta  
983 15 50.0 177 1 A55537 CDPdiacylglycerol-  
984 15 50.0 177 2 C71329 hypothetical prote  
985 15 50.0 178 2 T29345 hypothetical prote  
986 15 50.0 178 2 C84723 probable PHD-type  
987 15 50.0 178 2 D75390 NADH dehydrogenase  
988 15 50.0 180 2 S14019 hypothetical prote  
989 15 50.0 180 2 E84768 hypothetical prote  
990 15 50.0 180 2 F75607 arsenate reductase  
991 15 50.0 181 2 D86518 hypothetical prote  
992 15 50.0 181 2 C72104 hypothetical prote  
993 15 50.0 182 2 T45302 hypothetical prote  
994 15 50.0 182 2 T11906 NADH2 dehydrogenas  
995 15 50.0 183 2 A85561 hypothetical prote  
996 15 50.0 183 2 D64795 Apo-citrate lyase  
997 15 50.0 183 2 E90710 Apo-citrate lyase  
998 15 50.0 184 2 AG0736 hypothetical prote  
999 15 50.0 185 1 JC2394 phospholipase A2 1  
1000 15 50.0 185 2 C96808 protein F28K19.4 [

## ALIGNMENTS

## RESULT 1

Unknown protein, 30275-30808 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C/Accession: F96719  
R/Theologidis, A.; Eckert, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Comp, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Hutzler, L.  
Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Maritali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: F96719

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-177 <STO>

A/Cross-references: UNIPROT:P96C9L7; UNIPARC:UPI00000A1912; GB:AE005173; NID:96665540; PI

C/Genetics:

A/Gene: T6C23.4

A/Map position: 1

Query Match 60.0%; Score 18; DB 2; Length 177;  
Best Local Similarity 25.0%; Pred. No. 66;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 87 CSSSSSSC 94

## RESULT 2

A34597 rhomboid protein - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster  
C/Date: 06-Jul-1990 #sequence\_revision 06-Jul-1990 #text\_change 05-Oct-2004

C/Accession: A34597

R/Bier, E.; Jan, L.Y.; Jan, Y.N.

Genes Dev. 4, 190-203, 1990

A/Title: rhomboid, a gene required for dorsoventral axis establishment and peripheral ne

A/Reference number: A34597; MUID:90249726; PMID:2110920

A/Accession: A34597

A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-355 <BI>  
A/Cross-references: UNIPROT:P20350; UNIPARC:UPI000016BD27; GB:X52454; NID:910878; PIDN:C

C/Genetics:

A/Gene: FlyBase:ve

A/Cross-references: FlyBase:FBgn0004635

C/Superfamily: AAA protein

C/Keywords: transmembrane protein

Query Match 60.0%; Score 18; DB 2; Length 355;  
Best Local Similarity 25.0%; Pred. No. 82;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 57 CSTASTC 64

## RESULT 3

TPA-induced protein 11D - mouse

C/Species: Mus musculus (house mouse)  
C/Date: 30-Aug-1991 #sequence\_revision 30-Aug-1991 #text\_change 09-Jul-2004

C/Accession: C39590  
R/Varnum, B.C.; Ma, O.; Chi, T.; Fletcher, B.; Herschman, H.R.

Mol. Cell. Biol. 11, 1754-1758, 1991

A/Title: The Tis11 primary response gene is a member of a gene family that encodes prote

A/Reference number: A39590; MUID:91141531; PMID:1996120

A/Accession: C39590

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-367 <YAR>

A/Cross-references: UNIPROT:P23949; UNIPARC:UPI000002A130; GB:M58564; NID:9202066; PIDN

C/Keywords: DNA binding; zinc finger

Query Match 60.0%; Score 18; DB 2; Length 367;  
Best Local Similarity 25.0%; Pred. No. 82;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 273 CSSSASSC 280

## RESULT 4

S49147 ERF-2 protein - human

C/Species: Homo sapiens (man)  
C/Date: 16-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004

C/Accession: S49147; I37372  
R/Nie, X.F.; Maclean, R.N.; Kumar, V.; McKrady, I.A.; Bustin, S.A.

submitted to the EMBL Data Library, April 1994

A/Reference number: S49147

A/Accession: S49147

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-492 <NIE>

A/Cross-references: UNIPROT:P47974; UNIPARC:UPI000003794E

R/Nie, X.F.; Maclean, R.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.

Gene 152, 285-286, 1995

A/Title: ERF-2, the human homologue of the murine Tis11d early response gene.

A/Reference number: I37372; MUID:95137407; PMID:7835719

A/Accession: I37372

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-492 <RES>

A/Cross-references: UNIPARC:UPI000003794E; EMBL:X78992; NID:9509777; PIDN:CA55592.1; PI

C/Genetics:

A/Gene: GDB:BRF2; ERF2

A/Cross-references: GDB:511261

A/Map position: 13q12.3-13q12.3

Query Match 60.0%; Score 18; DB 2; Length 492;  
Best Local Similarity 25.0%; Pred. No. 90;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 299 CASSSSSC 306

## RESULT 5

JDVLJ1  
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adw, strain Japan/  
C/Species: hepatitis B virus, HBV  
C/Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 09-Jul-2004  
C/Accession: D28925  
R/Okamoto, H.; Tsuda, F.; Sakigawa, H.; Sastrosoewignjo, R.I.; Imai, M.; Miyakawa, Y.; W  
J. Gen. Virol. 69, 2573-2583, 1988  
A/Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of sur  
A/Reference number: J50253; MWID:89010694; PMID:3171552  
A/Accession: D28925  
A/Molecule type: DNA  
A/Residues: 1-843 <OKA>  
A/Cross-references: UNIPROT:P17394; UNIPARC:UPI00001297F9; GB:D00329; NID:G221497  
C/Superfamily: hepatitis virus DNA-directed DNA polymerase  
C/Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 60.0%; Score 18; DB 1; Length 843;  
Best Local Similarity 25.0%; Pred. No. 1.1e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 268 CASSSSSC 275

## RESULT 6

JDVLJ2  
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adw, strain Okinaw  
C/Species: hepatitis B virus, HBV  
C/Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 09-Jul-2004  
C/Accession: E28925  
R/Okamoto, H.; Tsuda, F.; Sakigawa, H.; Sastrosoewignjo, R.I.; Imai, M.; Miyakawa, Y.; W  
J. Gen. Virol. 69, 2573-2583, 1988  
A/Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of sur  
A/Reference number: J50253; MWID:89010694; PMID:3171552  
A/Accession: E28925  
A/Molecule type: DNA  
A/Residues: 1-843 <OKA>  
A/Cross-references: UNIPROT:P17395; UNIPARC:UPI00001297FC; GB:D00330; NID:G221498  
C/Superfamily: hepatitis virus DNA-directed DNA polymerase  
C/Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 60.0%; Score 18; DB 1; Length 843;  
Best Local Similarity 25.0%; Pred. No. 1.1e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 268 CASSSSSC 275

## RESULT 7

T26859  
hypothetical protein Y43F8B.3 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T26859  
R/Ainscough, R.  
submitted to the EMBL Data Library, October 1998  
A/Reference number: Z20278  
A/Accession: T26859  
A/Status: preliminary; translated from GB/EMBL/DBD  
A/Molecule type: DNA

A/Residues: 1-1743 <WIL>  
A/Cross-references: UNIPROT:Q9XMX5; UNIPARC:UPI0000078FD6; EMBL:AL032623; PIDN:CAA21511  
A/Experimental source: clone Y43F8B  
C/Genetics:  
A/Gene: CESP:Y43F8B.3  
A/Intons: 65/1, 92/2, 128/1, 229/1, 367/1, 422/1, 486/2, 523/1, 571/1, 628/1, 857/2, 9

Query Match 60.0%; Score 18; DB 2; Length 1743;  
Best Local Similarity 25.0%; Pred. No. 1.3e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 1123 CSATSTTC 1130

## RESULT 8

T31687  
surface antigen - Parametium primaurelia  
C/Species: Parametium primaurelia  
C/Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C/Accession: T31687  
R/Bourgain-Guglielmetti, F.; Caron, F.  
Journal of Eukaryot. Microbiol. 43, 303-314, 1996  
A/Title: Molecular characterization of the D surface protein gene subfamily in Paramet  
A/Reference number: Z21061; MWID:96313351; PMID:8768434  
A/Accession: T31687  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-2543 <BOU>  
A/Cross-references: UNIPROT:P90649; UNIPARC:UPI000007A903; EMBL:X96616; NID:G1235576; P  
C/Genetics:  
A/Genetic code: SGC5  
C/Superfamily: G surface protein

Query Match 60.0%; Score 18; DB 2; Length 2543;  
Best Local Similarity 25.0%; Pred. No. 1.5e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 1961 CSTSTTTC 1968

## RESULT 9

A44036  
collagen alpha 1(XII) chain - bovine (fragment)  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C/Accession: A44036  
R/Lunstrum, G.P.; McDonough, A.M.; Marinkovich, M.P.; Keene, D.R.; Morris, N.P.; Burges  
J. Biol. Chem. 267, 20087-20092, 1992  
A/Title: Identification and partial purification of a large, variant form of type XII c  
A/Reference number: A44036; MWID:93015874; PMID:1400326  
A/Accession: A44036  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-26 <LUN>  
A/Cross-references: UNIPROT:Q9TRM6; UNIPARC:UPI0000087EAB  
A/Experimental source: skin  
A/Note: sequence extracted from NCBI backbone (NCBIP:115795)

Query Match 56.7%; Score 17; DB 2; Length 26;  
Best Local Similarity 25.0%; Pred. No. 89;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 10 CETATSSC 17

## RESULT 10

S72742

B1177.F1.32 protein - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: S72742  
R:Smith, D.R.; Robison, K.  
Submitted to the EMBL Data Library, November 1993  
A:Description: Mycobacterium leprae cosmid B1177.  
A:Reference number: S72694  
A:Accession: S72742  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-89 <SMI>  
A:Cross-references: UNIPROT:Q49852; UNIPARC:UPI0000086610; EMBL:U00011; NID:9466807; PID:G6metics  
A:Start codon: GTG

Query Match 56.7%; Score 17; DB 2; Length 89;  
Best Local Similarity 25.0%; Pred. No. 1.3e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
| |  
DB 71 CRTSTRAC 78

RESULT 11  
T42262  
hypothetical protein - phage SPPI  
C:Species: phage SPPI  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T42262  
R:Alonso, J.C.; Luder, G.; Stiege, A.C.; Chai, S.; Weise, F.; Trautner, T.A.  
Gene 204, 201-212, 1997  
A:Title: The complete nucleotide sequence and functional organization of Bacillus subtilis  
A:Reference number: Z22137, MUID:98094274; PMID:9434185  
A:Accession: T42262  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-101 <MO>  
A:Cross-references: UNIPROT:Q48435; UNIPARC:UPI0000095544; EMBL:X97918; PIDN:CAA66572.1

Query Match 56.7%; Score 17; DB 2; Length 101;  
Best Local Similarity 25.0%; Pred. No. 1.3e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
| |  
DB 59 CSKTSAC 66

RESULT 12  
S72979  
probable L-a-aminoadipyl-L-cystenyl-D-valine synthetase acvs - Mycobacterium leprae  
N:Alternate names: B29 C2 179 protein  
C:Species: Mycobacterium leprae  
C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: S72979  
R:Smith, D.R.; Robison, K.  
Submitted to the EMBL Data Library, November 1993  
A:Description: Mycobacterium leprae cosmid B229.  
A:Reference number: S72588  
A:Accession: S72979  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-120 <SMI>  
A:Cross-references: UNIPROT:Q49859; UNIPARC:UPI00000862D7; EMBL:U00020; NID:9467102; PID:G6metics  
A:Gene: acvs  
A:Start codon: GTG

Query Match 56.7%; Score 17; DB 2; Length 120;  
Best Local Similarity 25.0%; Pred. No. 1.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
| |  
DB 4 CSYTSASC 11

RESULT 13  
JQ0150  
hypothetical 13k protein - Pseudomonas aeruginosa  
C:Species: Pseudomonas aeruginosa  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Sep-1999  
C:Accession: JQ0150  
R:Kato, J.; Chu, L.; Kitano, K.; Devault, J.D.; Kimura, K.; Chakrabarty, A.M.; Miera, T.  
Gene 84, 31-38, 1989  
A:Title: Nucleotide sequence of a regulatory region controlling alginate synthesis in Ps  
A:Reference number: JQ0132; MUID:90108714; PMID:2514124  
A:Accession: JQ0150  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-122 <KAT>  
A:Cross-references: UNIPARC:UPI00001791D0  
C:Superfamily: Pseudomonas aeruginosa hypothetical 13k protein

Query Match 56.7%; Score 17; DB 2; Length 122;  
Best Local Similarity 25.0%; Pred. No. 1.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
| |  
DB 49 CSTATVAC 56

RESULT 14  
S25574  
IG heavy chain V region - human  
C:Species: Homo sapiens (man)  
C:Date: 25-Feb-1994 #sequence\_revision 26-May-1995 #text\_change 20-Jun-2000  
C:Accession: S25574  
R:Addey, B.E.; Shackelford, P.G.; Quinn, A.; Carroll, W.L.  
J. Immunol. 147, 1667-1674, 1991  
A:Title: Restricted Ig H chain V gene usage in the human antibody response to Haemophilu  
A:Reference number: S25571; MUID:91349592; PMID:1908880  
A:Accession: S25574  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-123 <ADD>  
A:Cross-references: UNIPARC:UPI00001158BF; EMBL:X56527; NID:933569; PIDN:CAA39874.1; PID  
A:Note: the authors translated the codon GAA for residue 69 as Asp and TTA for residue 1  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:23-108/Domain: immunoglobulin homology <IMM>

Query Match 56.7%; Score 17; DB 2; Length 123;  
Best Local Similarity 25.0%; Pred. No. 1.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
| |  
DB 113 CASSAYAC 120

RESULT 15  
RHM0VY  
T-cell receptor beta chain precursor V region (Y135) - human  
N:Alternate names: T-cell receptor beta-1 chain J-B1.2 segment  
C:Species: Homo sapiens (man)  
C:Date: 03-Aug-1984 #sequence\_revision 03-Aug-1984 #text\_change 09-Jul-2004  
C:Accession: R02000; E24687  
R:Yanagi, Y.; Yoshikai, Y.; Leggett, K.; Clark, S.P.; Alexander, I.; Mak, T.W.  
Nature 308, 145-149, 1984  
A:Title: A human T cell-specific cDNA clone encodes a protein having extensive homology  
A:Reference number: A93324; MUID:84142269; PMID:6336315  
A:Accession: A02000



A/Molecule type: mRNA  
 A/Residues: 1-135 <YAN>  
 A/Cross-references: UNIPROT:P01733; UNIPARC:UPI00001377D3; GB:K01571  
 A/Experimental source: clone Y735  
 R/Toyonaga, B.; Yoshikai, Y.; Vadasz, V.; Chin, B.; Mak, T.W.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 8624-8628, 1985  
 A/Title: Organization and sequences of the diversity, joining, and constant region genes  
 A/Reference number: A94081; MUID:86094276; PMID:3866244  
 A/Accession: E24687  
 A/Molecule type: DNA  
 A/Residues: 121-135 <TOY>  
 A/Cross-references: UNIPARC:UPI0000113CT6; GB:M14158; NID:g338844; PIDN:AAA60669.1; PID:  
 C/Genetics:  
 A/Status: GDB:TCRB  
 A/Cross-references: GDB:120405; OMIM:186930  
 A/Map position: 7q35-7q35  
 C/Superfamily: Immunoglobulin V region; immunoglobulin homology  
 C/Keywords: glycoprotein; heterotrimer; receptor; T-cell  
 F/1-21/Domain: signal sequence #status predicted <SIG>  
 F/22-135/Product: T-cell receptor beta chain V region Y735 #status predicted <MAT>  
 F/35-113/Domain: immunoglobulin homology <IMM>  
 F/42-111/Disulfide bonds: #status predicted  
 F/90/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 56.7%; Score 17; DB 1; Length 135;  
 Best Local Similarity 25.0%; Pred. No. 1.5e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 111 CASSFSTC 118

RESULT 16  
 A71123  
 hypothetical protein PH0752 - Pyrococcus horikoshii  
 C/Species: Pyrococcus horikoshii  
 C/Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 12-Jul-2004  
 C/Accession: A71123  
 R/Ramabadyaai, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
 M.; Ohtoku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
 DNA Res. 5, 55-76, 1998  
 A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A/Reference number: A71000; MUID:98344137; PMID:9679194  
 A/Accession: A71123  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-139 <RAM>  
 A/Cross-references: UNIPROT:O58508; UNIPARC:UPI0000062ECE; GB:AP000003; NID:g3236130; PI  
 A/Experimental source: strain OT3  
 A/Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C/Genetics:  
 A/Status: GDB:PH0752

Query Match 56.7%; Score 17; DB 2; Length 139;  
 Best Local Similarity 25.0%; Pred. No. 1.5e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 81 CSSTVASC 88

RESULT 17  
 H87436  
 hypothetical protein CC1513 [imported] - Caulobacter crescentus  
 C/Species: Caulobacter crescentus  
 C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
 C/Accession: H87436  
 R/Hierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gattm, M.L.; Haft, D.H.; Kolon  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of Caulobacter crescentus.  
 A/Reference number: A87249; MUID:21173698; PMID:11259647  
 A/Accession: H87436  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-145 <STO>  
 A/Cross-references: UNIPROT:Q9A851; UNIPARC:UPI00000C73F0; GB:AE005673; NID:g13422892;  
 C/Genetics:  
 A/Status: GDB:CC1513

Query Match 56.7%; Score 17; DB 2; Length 145;  
 Best Local Similarity 25.0%; Pred. No. 1.5e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 10 CAAABAC 17

RESULT 18  
 T22305  
 hypothetical protein F46C5.2 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T22305  
 R/Thomas, K.  
 submitted to the EMBL Data Library, September 1995  
 A/Reference number: Z19544  
 A/Accession: T22305  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-182 <WIL>  
 A/Cross-references: UNIPROT:P52881; UNIPARC:UPI000013A065; EMBL:Z54281; PIDN:CAA91050.1  
 A/Experimental source: clone F46C5  
 C/Genetics:  
 A/Status: GDB:F46C5.2  
 A/Map position: 2  
 A/Introns: 14/3; 50/1; 76/3; 145/2

Query Match 56.7%; Score 17; DB 2; Length 182;  
 Best Local Similarity 25.0%; Pred. No. 1.6e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 35 CTSGATTC 42

RESULT 19  
 E70641  
 hypothetical protein RV0698 - Mycobacterium tuberculosis (strain H37RV)  
 C/Species: Mycobacterium tuberculosis  
 C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
 C/Accession: E70641  
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
 ; Comor, R.; Davies, R.; Devlin, K.; Feltywell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A/Reference number: A70500; MUID:98295987; PMID:9634230  
 A/Accession: E70641  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-203 <COL>  
 A/Cross-references: UNIPROT:P95044; UNIPARC:UPI00000C14B5; GB:Z84395; GB:AL123456; NID:  
 A/Experimental source: strain H37RV  
 C/Genetics:  
 A/Status: GDB:RV0698

Query Match 56.7%; Score 17; DB 2; Length 203;  
 Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
|  
Db 51 CTTAHTC 58

RESULT 20  
H75277  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #ext\_change 09-Jul-2004  
C:Accession: H75277  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; PMID:20036896; PMID:10567266  
A:Accession: H75277  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-212 <WHI>  
A:Cross-references: UNIPROT:Q9R88; UNIPARC:UPI00000CIA7; GB:AE002071; GB:AE000513; NID  
C:Genetics:  
A:Gene: DR2407  
A:Map position: 1

Query Match 56.7%; Score 17; DB 2; Length 212;  
Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
|  
Db 15 CRSASAC 22

RESULT 21  
A41128  
myogenin - human  
N:Alternate names: myogenic determination factor Myf-4  
C:Species: Homo sapiens (man)  
C>Date: 03-Aug-1992 #sequence\_revision 03-Aug-1992 #ext\_change 09-Jul-2004  
C:Accession: A41128; S27295; S06948; S21978  
R:Salminen, A.; Branu, T.; Buchberger, A.; Ueters, S.; Winter, B.; Arnold, H.H.  
J. Cell Biol. 115, 905-917, 1991  
A:Title: Transcription of the muscle regulatory gene MYF4 is regulated by serum component  
A:Reference number: A41128; PMID:92064650; PMID:1659574  
A:Accession: A41128  
A:Molecule type: DNA  
A:Residues: 1-224 <SAL>  
A:Cross-references: UNIPROT:P15173; UNIPARC:UPI000012FBA4; EMBL:X62155  
R:Branu, T.; Bober, E.; Buschhausen-Denker, G.; Koltz, S.; Grzeschik, K.H.; Arnold, H.H.  
EMBO J. 9, 592, 1990  
A:Reference number: S27295  
A:Contents: cistron  
A:Accession: S27295  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-224 <BRA>  
A:Cross-references: UNIPARC:UPI000012FBA4  
A:Note: This is a revision to the sequence from reference S06947  
R:Branu, T.; Bober, E.; Buschhausen-Denker, G.; Koltz, S.; Grzeschik, K.H.; Arnold, H.H.  
EMBO J. 8, 3617-3625, 1989  
A:Title: Differential expression of myogenic determination genes in muscle cells: possib  
A:Reference number: S06947; PMID:90059960; PMID:2583111  
A:Accession: S06948  
A:Molecule type: mRNA  
A:Residues: 1-154, S06948A  
A:Cross-references: UNIPARC:UPI000016AD95; EMBL:X17651; NID:934831; PIDN:CAA5641.1; PID  
A:Note: this sequence has been revised in reference S27295  
R:Arnold, H.H.  
submitted to the EMBL Data Library, September 1991

A:Reference number: S21978  
A:Accession: S21978  
A:Molecule type: DNA  
A:Residues: 1-223 <ARN>  
A:Cross-references: UNIPARC:UPI000016AD96; EMBL:X62155; NID:934833; PIDN:CAA44080.1; PID  
C:Genetics:  
A:Gene: GDB:MYOG; MYF4  
A:Cross-references: GDB:120210; OMIM:159980  
A:Map position: 1q31-1q41  
A:Introns: 157/3; 184/3  
C:Superfamily: human myogenin  
C:Keywords: DNA binding

Query Match 56.7%; Score 17; DB 2; Length 224;  
Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
|  
Db 162 CSHASAC 169

RESULT 22  
S09778  
hypothetical protein UL16 precursor - human cytomegalovirus (strain AD169)  
C:Species: human cytomegalovirus, human herpesvirus 5  
A:Note: host Homo sapiens (man)  
C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #ext\_change 09-Jul-2004  
C:Accession: S09778  
R:Chen, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;  
M.; Barrett, B.G.  
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990  
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus  
A:Reference number: S09749; PMID:90269039; PMID:2161319  
A:Accession: S09778  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-230 <CHE>  
A:Cross-references: UNIPROT:P16757; UNIPARC:UPI000017B56; EMBL:X17403; NID:959591; PIDN  
A:Note: this sequence was submitted to the EMBL Data Library, December 1989  
C:Superfamily: human cytomegalovirus hypothetical protein UL16  
C:Keywords: glycoprotein; transmembrane protein  
F:1-25/Domain: signal sequence #status predicted <SIG>  
F:26-230/Product: hypothetical protein UL16 #status predicted <MAT>  
F:189-206/Domain: transmembrane #status predicted <TM>  
F:35,41,68,84,95,101,132,145/Binding site: carbohydrate (Aem) (covalent) #status predict

Query Match 56.7%; Score 17; DB 2; Length 230;  
Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
|  
Db 18 CUSASAC 25

RESULT 23  
P95411  
hypothetical protein Sma221 [imported] - Sinorhizobium meliloti (strain 1021) magaplae  
C:Species: Sinorhizobium meliloti  
C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #ext\_change 09-Jul-2004  
C:Accession: P95411  
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow  
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9889, 2001  
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo  
A:Reference number: A95262; PMID:21396509; PMID:11481432  
A:Accession: P95411  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-237 <MUR>  
A:Cross-references: UNIPROT:Q92X56; UNIPARC:UPI00000CB344; GB:AE006469; PIDN:AAK65856.1  
A:Experimental source: strain 1021, megaplasmid pSymA

R:Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, J.; D.; Chain, P.; Cowie, A.; Davies, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, J.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Welle, D.H.; Wong, K.; Yen, K.  
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: Sma2221  
 A:Genome: plasmid

Query Match 56.7%; Score 17; DB 2; Length 237;  
 Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 76 CTAITTC 83

## RESULT 24

JH0367  
 ribonuclease (EC 3.1.27.-) Rh precursor - *Rhizopus niveus*  
 N:Alternate names: base nonspecific ribonuclease  
 C:Species: *Rhizopus niveus*  
 C>Date: 05-Mar-1993 #sequence\_revision 31-Jan-1997 #text\_change 31-Dec-2004  
 C/Accession: JH0367; A41449; B61449  
 R:Ohg1, K.; Horluch, H.; Watanabe, H.; Takagi, M.; Yano, K.; Irie, M.  
 J. Biochem. 109, 776-785, 1991  
 A:Title: Expression of RNase R from *Rhizopus niveus* in yeast and characterization of th  
 A:Reference number: JH0367; MUID:92011465; PMID:1655721  
 A:Accession: JH0367  
 A:Molecule type: DNA  
 A:Residues: 1-238 <OHG>  
 A:Cross-references: UNIPROT:P08056; UNIPARC:UPI00001344AC; DDBJ:D12476; DDBJ:D01125; NID  
 R:Horluch, H.; Yano, K.; Takagi, M.; Yano, K.; Wakabayashi, E.; Sanda, A.; Mine, S.; C  
 J. Biochem. 103, 408-418, 1988  
 A:Title: Primary structure of a base non-specific ribonuclease from *Rhizopus niveus*.  
 A:Reference number: A41449; MUID:88273061; PMID:3391995  
 A:Accession: A41449  
 A:Molecule type: DNA  
 A:Residues: 1-81, 'STY', 85-238 <HOR>  
 A:Cross-references: UNIPARC:UPI000016897A; GB:D00238; NID:G218044; PIDN:BAA00167.1; PID  
 A:Note: the authors translated the codon TCA for residue 82 as Asn, CTC for residue 83 a  
 A:Accession: B41449  
 A:Molecule type: protein  
 A:Residues: 17-238 <HO2>  
 A:Cross-references: UNIPARC:UPI00001109AA  
 R:Kurthar, H.; Nonaka, T.; Mitsui, Y.; Ohg1, K.; Irie, M.; Nakamura, K.T.  
 J. Mol. Biol. 255, 310-320, 1996  
 A:Title: The crystal structure of ribonuclease Rh from *Rhizopus niveus* at 2.0 Å resolution  
 A:Reference number: S62310; MUID:96150305; PMID:851522  
 A:Contents: annotation; X-ray crystallography, 2.0 angstroms  
 C:Genetics:  
 A:Intons: 21/1; 116/2; 151/3; 167/2  
 C:Function:  
 A:Description: hydrolyzes internal phosphodiester bonds of RNA to produce 3'-phosphomono  
 C:Superfamily: RNases  
 C:Keywords: hydrolase  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:17-238/Product: ribonuclease Rh #status experimental <MAT>  
 F:19-36,26-69,35-136,79-128,198-229/Disulfide bonds: #status experimental  
 F:52,121,125/Active site: His, Gln, His #status predicted  
 F:65,120,124/Binding site: substrate (Trp, His, Lys) #status predicted

Query Match 56.7%; Score 17; DB 1; Length 238;  
 Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

DB 19 CSSTALSC 26

## RESULT 25

A35871  
 trypsin (EC 3.4.21.4) precursor, pancreatic - African clawed frog  
 C:Species: *Xenopus laevis* (African clawed frog)  
 C>Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 09-Jul-2004  
 C/Accession: A35871; S12117  
 R:Shi, Y.B.; Brown, D.D.  
 Genes Dev. 4, 1107-1113, 1990  
 A:Title: Developmental and thyroid hormone-dependent regulation of pancreatic genes in  
 A:Reference number: A35871; MUID:91007255; PMID:2210372  
 A:Accession: A35871  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-243 <SHI>  
 A:Cross-references: UNIPROT:P19799; UNIPARC:UPI00001376B0; EMBL:X53458; NID:965162; PID  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; protein digestion; serine proteinase  
 F:1-15/Domain: signal sequence #status predicted <SIG>  
 F:16-20/Domain: activation peptide #status predicted <APR>  
 F:21-236/Domain: trypsin homology <TRY>  
 F:26-243/Product: trypsin I #status predicted <MAT>  
 F:27-157,45-61,129-230,136-203,168-182/Disulfide bonds: #status predicted  
 F:60,104,197/Active site: His, Asp, Ser #status predicted

Query Match 56.7%; Score 17; DB 2; Length 243;  
 Best Local Similarity 25.0%; Pred. No. 1.8e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 129 CSAAGTSC 136

## RESULT 26

T47902  
 hypothetical protein T4C21.320 - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
 C/Accession: T47902  
 R:Choiane, N.; Robert, C.; Brothier, P.; Wincker, P.; Catolico, L.; Artiguenave, F.; S.  
 W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.  
 submitted to the Protein Sequence Database, March 2000  
 A:Reference number: Z24479  
 A:Accession: T47902  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-248 <CHO>  
 A:Cross-references: UNIPROT:Q9LZX3; UNIPARC:UPI00009C7A5; EMBL:AL162295  
 A:Experimental source: cultivar Columbia; BAC clone T4C21  
 C:Genetics:  
 A:Map position: 3  
 A:Intons: 61/3; 96/1; 119/1; 125/3; 143/2; 153/3; 181/1  
 A:Note: T4C21.320

Query Match 56.7%; Score 17; DB 2; Length 248;  
 Best Local Similarity 25.0%; Pred. No. 1.8e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 191 CSSAKSC 198

## RESULT 27

E87269  
 exopolysaccharide production protein Pse [imported] - *Caulobacter crescentus*  
 C:Species: *Caulobacter crescentus*  
 C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 31-Dec-2004  
 C/Accession: E87269  
 R:Meriman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.B.; Eiseen, J.; Heidelberg, J

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Kholmova, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; UID:21173698; PMID:11259647

A:Accession: E87269  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-267 <STO>  
A:Cross-references: UNIPROT:Q9ABR0; UNIPARC:UPI000006CF59; GB:AE005673; NID:913421281; F  
C:Genetics:  
A:Gene: CC0166  
C:Superfamily: sugar transferase, ExoY type

Query Match 56.7%; Score 17; DB 2; Length 267;  
Best Local Similarity 25.0%; Pred. No. 1.8e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 20 CTTTSSDC 27

RESULT 28  
F70771  
probable glutamateracemase - *Mycobacterium tuberculosis* (strain H37RV)  
C:Species: *Mycobacterium tuberculosis*  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: F70771

R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holtrope, S.  
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
A:Reference number: A70500; PMID:98295987; PMID:9634230

A:Accession: F70771  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-271 <COL>  
A:Cross-references: UNIPROT:Q10626; UNIPARC:UPI000012PA0B; GB:Z73902; GB:AL123456; NID:9  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: mtrI  
C:Superfamily: glutamate racemase

Query Match 56.7%; Score 17; DB 2; Length 271;  
Best Local Similarity 25.0%; Pred. No. 1.8e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 75 CNGSSSAC 82

RESULT 29  
S72790  
glutamate racemase (EC 5.1.1.3) - *Mycobacterium leprae*  
N:Alternate names: B1549 C2 210 protein  
C:Species: *Mycobacterium leprae*  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: S72790  
R:Smith, D.R.; Robison, K.  
submitted to the EMBL Data Library, November 1993  
A:Description: *Mycobacterium leprae* cosmid B1549.  
A:Reference number: S72582  
A:Accession: S72790  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-272 <SMI>  
A:Cross-references: UNIPROT:P46705; UNIPARC:UPI000012PA0D; EMBL:U00014; NID:9466903; PTD  
C:Genetics:  
A:Gene: mtrI

C:Superfamily: glutamate racemase  
C:Keywords: Isomerase

Query Match 56.7%; Score 17; DB 1; Length 272;  
Best Local Similarity 25.0%; Pred. No. 1.8e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 75 CMTASAC 82

RESULT 30  
I51172  
transcription factor RCC/BBP-2 - bullfrog  
C:Species: *Rana catesbeiana* (bullfrog)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C:Accession: I51172  
R:Chen, Y.; Hu, H.; Atkinson, B.G.  
Dev. Genet. 15, 366-377, 1994  
A:Title: Characterization and expression of C/EBP-like genes in the liver of *Rana catesbeiana*  
A:Reference number: I51171; UID:95008407; PMID:7923939  
A:Accession: I51172  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-284 <CHE>  
A:Cross-references: UNIPROT:Q91295; UNIPARC:UPI000007BACC; EMBL:U08605; NID:9478890; PTD  
C:Superfamily: CCMAT/enhancer-binding protein alpha

Query Match 56.7%; Score 17; DB 2; Length 284;  
Best Local Similarity 25.0%; Pred. No. 1.8e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 171 CSNTSSAC 178

RESULT 31  
C64795  
citG protein - *Escherichia coli* (strain K-12)  
C:Species: *Escherichia coli*  
C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
C:Accession: C64795  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of *Escherichia coli* K-12.  
A:Reference number: A64720; UID:97426617; PMID:9278503  
A:Accession: C64795  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-292 <BLAT>  
A:Cross-references: UNIPROT:P77231; UNIPARC:UPI0000127A40; GB:AE000166; GB:U00096; NID:9  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: citG  
C:Superfamily: citG protein

Query Match 56.7%; Score 17; DB 2; Length 292;  
Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 142 CSTASFC 149

RESULT 32  
D90710  
2-(5'-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase (EC 4.2.-.-) EC0652 - *Escheri*  
C:Species: *Escherichia coli*  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004

C:Accession: D90710  
 R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
 Sasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno  
 A:Reference number: A9629; UID:2115631; PMID:11258796  
 A:Accession: D90710  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-292 <HAY>  
 A:Cross-references: UNIPROT:P58161; UNIPARC:UPI0000127A3F; GB:BA000007; PIDN:BA834075.1;  
 A:Experimental source: strain O157:H7, substrain RMD 0509952  
 C:Genetics:  
 A:Gene: BCS0652  
 C:Superfamily: citg protein  
 C:Keywords: carbon-oxygen lyase

Query Match 56.7%; Score 17; DB 2; Length 292;  
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 142 CSTAASFC 149

RESULT 33  
 H85560  
 2-(5'-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase (EC 4.2.-.-) citg - *Escherichia*  
 C:Species: *Escherichia coli*  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: H85560  
 R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; DiMantola, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A:Reference number: A85480; UID:21074935; PMID:11206551  
 A:Accession: H85560  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-292 <STO>  
 A:Cross-references: UNIPROT:P58161; UNIPARC:UPI0000127A3F; GB:AE005174; NID:G12513512; F  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: citg  
 C:Superfamily: citg protein  
 C:Keywords: carbon-oxygen lyase

Query Match 56.7%; Score 17; DB 2; Length 292;  
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 142 CSTAASFC 149

RESULT 34  
 A98294  
 Hypothetical protein AGR\_L\_2619 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere  
 C:Species: *Agrobacterium tumefaciens*  
 C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
 C:Accession: A98294  
 R.Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Mollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2333-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*  
 A:Reference number: A97359; UID:21608551; PMID:11743194  
 A:Accession: A98294  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-297 <KUR>  
 A:Cross-references: UNIPROT:O8U53; UNIPARC:UPI00000D222C; GB:AE007870; PIDN:AAK89875.1;

C:Genetics:  
 A:Gene: AGR\_L\_2619  
 A:Map position: linear chromosome

Query Match 56.7%; Score 17; DB 2; Length 297;  
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 151 CAAVAAC 158

RESULT 35  
 AH2989  
 permease [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)  
 C:Species: *Agrobacterium tumefaciens*  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
 C:Accession: AH2989  
 R.Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavich, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
 A:Reference number: AB2577; UID:21608550; PMID:11743193  
 A:Accession: AH2989  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-297 <KUR>  
 A:Cross-references: UNIPROT:O8U53; UNIPARC:UPI00000D222C; GB:AE008689; PIDN:AL44334.1,  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: ACU3521  
 A:Map position: linear chromosome

Query Match 56.7%; Score 17; DB 2; Length 297;  
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 151 CAAVAAC 158

RESULT 36  
 AE0857  
 conserved hypothetical protein STY3065 [imported] - *Salmonella enterica* subsp. *enterica*  
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh  
 A:Note: this species has also been called *Salmonella typhi*  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C:Accession: AE0857  
 R.Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
 , S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.,  
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov  
 A:Reference number: AB0502; UID:21534947; PMID:11677608  
 A:Accession: AE0857  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-305 <PAR>  
 A:Cross-references: UNIPARC:UPI000005A39C; GB:AL513382; PIDN:CAD06044.1; PID:G16504011;  
 C:Genetics:  
 A:Gene: STY3065  
 C:Superfamily: *Escherichia coli* hypothetical protein b2755

Query Match 56.7%; Score 17; DB 2; Length 305;  
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
|  
Db 179 CISAATSC 186

## RESULT 37

GS5056  
hypothetical protein b2755 - Escherichia coli (strain K-12)  
C/Species: Escherichia coli  
C/Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
C/Accession: G65056  
R/Biotech: F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A/Title: The complete genome sequence of Escherichia coli K-12.  
A/Reference number: A64720; MUID:97426617; PMID:9278503  
A/Accession: G65056  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-305 <BLAT>  
A/Cross-references: UNIPROT:Q46896; UNIPARC:UPI000013B034; GB:AE000359; GB:U00096; NID:9  
A/Experimental source: strain K-12, substrain MG1655  
C/Superfamily: Escherichia coli hypothetical protein b2755

Query Match 56.7%; Score 17; DB 2; Length 305;  
Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
|  
Db 179 CISAATSC 186

## RESULT 38

B85925  
hypothetical protein Z4064 [imported] - Escherichia coli (strain O157:H7, substrain EDL9  
C/Species: Escherichia coli  
C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: B85925  
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A/Reference number: A85480; MUID:21074935; PMID:11206551  
A/Accession: B85925  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-307 <STO>  
A/Cross-references: UNIPROT:Q8X7M3; UNIPARC:UPI000000D09E4; GB:AE005174; NID:912517212; F  
A/Experimental source: strain O157:H7, substrain EDL933  
C/Genetics:  
A/Gene: Z4064  
C/Superfamily: Escherichia coli hypothetical protein b2755

Query Match 56.7%; Score 17; DB 2; Length 307;  
Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
|  
Db 179 CISAATSC 186

## RESULT 39

A91080  
hypothetical protein EC03609 [imported] - Escherichia coli (strain O157:H7, substrain R1  
C/Species: Escherichia coli  
C/Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C/Accession: A91080  
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasegawa, N.; Yasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend

A/Reference number: A99629; MUID:21156231; PMID:11258796

A/Accession: A91080  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-307 <HAV>  
A/Cross-references: UNIPROT:Q8X7M3; UNIPARC:UPI000000D09E4; GB:BA000007; PIDN:BAAB37032.1;  
A/Experimental source: strain O157:H7, substrain R1MD 0509952  
C/Genetics:  
A/Gene: EC03609  
C/Superfamily: Escherichia coli hypothetical protein b2755

Query Match 56.7%; Score 17; DB 2; Length 307;  
Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
|  
Db 179 CISAATSC 186

## RESULT 40

T32376  
hypothetical protein K10F12.4 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T32376  
R/Wohlmann, P.; Beck, C.  
submitted to the EMBL Data Library, September 1997  
A/Description: The sequence of C. elegans cosmid K10F12.  
A/Reference number: Z21157  
A/Accession: T32376  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-309 <MOH>  
A/Cross-references: UNIPROT:O17234; UNIPARC:UPI000007F612; EMBL:AF025462; PIDN:AA871002.  
A/Experimental source: strain Bristol N2; clone K10F12  
C/Genetics:  
A/Gene: CESP:K10F12.4  
A/Map position: 3  
A/Introns: 31/3; 123/2; 196/3; 239/1

Query Match 56.7%; Score 17; DB 2; Length 309;  
Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
|  
Db 7 CSTSALC 14

## RESULT 41

B53522  
20k cyclophilin - Toxoplasma gondii (fragment)  
C/Species: Toxoplasma gondii  
C/Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C/Accession: B53522  
R/Hign, K.P.; Joiner, K.A.; Handchumacher, R.E.  
J. Biol. Chem. 269, 9105-9112, 1994  
A/Title: Isolation, cDNA sequences, and biochemical characterization of the major cyclo  
A/Reference number: A53522; MUID:94179329; PMID:8132648  
A/Accession: B53522  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-347 <HIG>  
A/Cross-references: UNIPROT:Q26995; UNIPARC:UPI000007DB6A; GB:U04634; NID:9436958; PID:  
F:178-346/Domain: cyclophilin homology <CYP>

Query Match 56.7%; Score 17; DB 2; Length 347;  
Best Local Similarity 25.0%; Pred. No. 2e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
|

Db 48 CASAEAC 55

RESULT 42

T36890

probable iron-siderophore uptake system transmembrane component - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C/Accession: T36890

R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A/Reference number: Z21617

A/Accession: T36890

A/Status: preliminary; translated from GB/EMBL/DDB

A/Molecule type: DNA

A/Residues: 1-348 <MUR>

A/Cross-references: UNIPROT:Q9S213; UNIPARC:UPI00000DB327; EMBL:AL109848; PIDN:CA852851.

A/Experimental source: strain A3(2)

C/Genetics:

A/Genes: SCOEDB:SC151.27c

C/Superfamily: vitamin B12 transport protein bnc

Query Match

Best Local Similarity 56.7%; Score 17; DB 2; Length 348;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

Db 262 CGAATPAC 269

RESULT 43

C84548

hypothetical protein At2g17120 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C/Accession: C84548

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, P.; VanAken, S.E.; Umayam, L.; Tallon, L.

Neus, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: C84548

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-350 <STO>

A/Cross-references: UNIPROT:O23006; UNIPARC:UPI00000A22D; GB:AE002093; NID:92558660; PI

C/Genetics:

A/Genes: At2g17120

A/Map position: 2

Query Match

Best Local Similarity 56.7%; Score 17; DB 2; Length 350;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

Db 31 CCGSTSTC 38

RESULT 44

B87507

conserved hypothetical protein CC2083 [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus

C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004

C/Accession: B87507

R:Nierman, W.C.; Felldbylum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DebRoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gilm, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of Caulobacter crescentus.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: B87507  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-352 <STO>  
A/Cross-references: UNIPROT:Q9A6U0; UNIPARC:UPI00000C75D3; GB:AE005673; NID:913423564;  
C/Genetics:  
A/Genes: CC2083

Query Match 56.7%; Score 17; DB 2; Length 352;

Best Local Similarity 25.0%; Pred. No. 2e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

Db 11 CAALSAC 18

RESULT 45

G71287

conserved hypothetical protein TP0730 - syphilis spirochete

C/Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C/Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004

C/Accession: G71287

R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwi

rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uteerback, T.; McD

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A/Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A/Reference number: A71250; MUID:98332770; PMID:9665876

A/Accession: G71287

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-365 <COV>

A/Cross-references: UNIPROT:O83712; UNIPARC:UPI00000C0AA2; GB:AE001245; GB:AE000520; NIT

A/Experimental source: strain Nichols

C/Genetics:

A/Genes: TP0730

Query Match

Best Local Similarity 56.7%; Score 17; DB 2; Length 365;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

Db 162 CAASDC 169

RESULT 46

D97175

nifs family enzyme (cysteine desulfurase/cysteine sulfinate desulfurase) [imported] - Cl

C/Species: Clostridium acetobutylicum 14-Sep-2001 #text\_change 09-Jul-2004

C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004

C/Accession: D97175

R:Molling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

.J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A/Reference number: A96900; MUID:21359325; PMID:21359325

A/Accession: D97175

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-376 <KUR>

A/Cross-references: UNIPROT:Q97GY1; UNIPARC:UPI00000D7546; GB:AE001437; PIDN:AAK80191.1,

A/Experimental source: Clostridium acetobutylicum ATCC824

C/Genetics:

A/Genes: CAC2234

C/Superfamily: nitrogen fixation protein nifs

Query Match 56.7%; Score 17; DB 2; Length 376;

Best Local Similarity 25.0%; Pred. No. 2e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

Db 319 CASGSGAC 326

## RESULT 47

C34443 nitrogenase cofactor synthesis protein nifS - Anabaena sp.

N:Contains: L-cysteine sulfotransferase (EC 2.8.1.-)

C:Species: Anabaena sp.

C:Date: 15-Jun-1990 #sequence\_revision 15-Jun-1990 #text\_change 09-Jul-2004

C:Accession: C34443; PMID:832361

R:Miller, M.E.; Haselkorn, R.

J. Biol. Chem. 264, 19200-19207, 1989

A:Title: Nitrogen fixation (nif) genes of the cyanobacterium Anabaena species strain PCC

A:Reference number: A34443; PMID:90037054; PMID:2553733

A:Accession: C34443

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-400 <MUL>

A:Cross-references: UNIPROT:P12623; UNIPARC:UPI000016E55C; GB:J05111; NID:g142034; PIDN:

R:Miller, M.E.; Bulikema, W.J.; Haselkorn, R.

J. Bacteriol. 170, 4405-4410, 1988

A:Title: Bacterial-type ferredoxin genes in the nitrogen fixation regions of the cyanoba

A:Reference number: A31884; PMID:86314954; PMID:2842320

A:Accession: B32361

A:Molecule type: DNA

A:Residues: 1-7 <MU2>

A:Cross-references: UNIPARC:UPI0000178A6F

C:Superfamily: nitrogen fixation protein nifS

C:Keywords: phosphoprotein, pyridoxal phosphate, sulfotransferase

F:201/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

F:324/Active site: Cys (cysteine persulfide intermediate) #status predicted

Query Match 56.7%; Score 17; DB 2; Length 400;

Best Local Similarity 25.0%; Pred. No. 2e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

Db 317 CASGSGAC 324

## RESULT 48

A31988 nitrogenase cofactor synthesis protein [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004

C:Accession: A31988

R:Kanehisa, T.; Nakamura, Y.; Wolik, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: A31807; PMID:2155285; PMID:11759840

A:Accession: A31988

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-400 <KUR>

A:Cross-references: UNIPROT:P12623; UNIPARC:UPI000013010C; GB:BA000019; PIDN:BAE73413.1;

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: nifS

C:Superfamily: nitrogen fixation protein nifS

Query Match 56.7%; Score 17; DB 2; Length 400;

Best Local Similarity 25.0%; Pred. No. 2e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

Db 317 CASGSGAC 324

## RESULT 49

MMNVIA

Immediate-early protein IE-N - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004

C:Accession: A39150

R:Carson, D.D.; Summers, M.D.; Guarino, L.A.

Virology 182, 279-286, 1991

A:Title: Molecular analysis of a baculovirus regulatory gene.

A:Reference number: A39150; PMID:91220660; PMID:2024466

A:Accession: A39150

A:Molecule type: DNA

A:Residues: 1-408 <CAR>

A:Cross-references: UNIPROT:P24647; UNIPARC:UPI0000170DEB; GB:M59422; NID:g332437; PIDN:

C:Superfamily: AcMNPV immediate-early protein IE-N

C:Keywords: DNA binding; early protein; tandem repeat; transcription regulation

F:34-49/Region: 7-residue repeats

F:51-58/Region: 4-residue repeats

F:190-196/Region: glutamine-rich

Query Match 56.7%; Score 17; DB 1; Length 408;

Best Local Similarity 25.0%; Pred. No. 2.1e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

Db 253 CSATSSDC 260

## RESULT 50

A72869 early gene transactivator - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV

A:Note: dsDNA virus

C:Date: 12-Nov-1999 #sequence\_revision 12-Nov-1999 #text\_change 09-Jul-2004

C:Accession: A72869

R:Byres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.

Virology 202, 586-605, 1994

A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.

A:Reference number: A72850; PMID:94303173; PMID:8030224

A:Accession: A72869

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-408 <AYR>

A:Cross-references: UNIPROT:P24647; UNIPARC:UPI00001387E9; GB:L22858; NID:g510708; PIDN:

C:Genetics:

A:Gene: Ac-IE-2

C:Superfamily: AcMNPV immediate-early protein IE-N

Query Match 56.7%; Score 17; DB 2; Length 408;

Best Local Similarity 25.0%; Pred. No. 2.1e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11

Db 253 CSATSSDC 260

Search completed: January 4, 2006, 16:09:57  
Job time : 28.487 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 91.4261 Seconds  
(without alignments)

108.037 Million cell updates/sec

Title: US-09-932-322-2  
Perfect score: 30  
Sequence: 1 XXXXXXXXXXXXXXX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database : Uniprot\_05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	60.0	69	2	Q9N669_GCAEN	Q9N669 conus ebraeae
2	18	60.0	69	2	Q9NCS8_GCAEN	Q9NCS8 conus ebraeae
3	18	60.0	69	2	Q9NCS9_GCAEN	Q9NCS9 conus ebraeae
4	18	60.0	69	2	Q9NCT2_GCAEN	Q9NCT2 conus ebraeae
5	18	60.0	69	2	Q9NCT3_GCAEN	Q9NCT3 conus ebraeae
6	18	60.0	69	2	Q9VD36_DROME	Q9VD36 drosophila
7	18	60.0	90	2	Q7QPP9_ANOGA	Q7QPP9 anophelies g
8	18	60.0	90	2	Q9B111_ANOGA	Q9B111 anophelies g
9	18	60.0	114	2	Q82854_GRETR	Q82854 jembrana di
10	18	60.0	161	2	Q9E0L8_MOUSE	Q9E0L8 mus musculu
11	18	60.0	163	2	Q7QPM4_GIALA	Q7QPM4 giardia lam
12	18	60.0	177	2	Q8LDT9_ARATH	Q8LDT9 arabidopsis
13	18	60.0	177	2	Q9C9L7_ARATH	Q9C9L7 arabidopsis
14	18	60.0	199	2	Q60M47_CABER	Q60M47 caenorhabdi
15	18	60.0	200	2	Q5VMP6_ORYSA	Q5VMP6 oryza sativ
16	18	60.0	224	2	Q9E0L7_MOUSE	Q9E0L7 mus musculu
17	18	60.0	233	2	Q6K2G5_ORYSA	Q6K2G5 oryza sativ
18	18	60.0	242	2	Q51XD3_MAGGR	Q51XD3 magnaporthe
19	18	60.0	245	2	Q6BCP3_HPBVO	Q6BCP3 hepatitis b
20	18	60.0	258	2	Q9Y0E9_DROME	Q9Y0E9 drosophila
21	18	60.0	269	2	Q7PR07_ANOGA	Q7PR07 anophelies g
22	18	60.0	319	2	Q7Y198_ORYSA	Q7Y198 oryza sativ
23	18	60.0	320	2	Q61AUS_CABER	Q61AUS caenorhabdi
24	18	60.0	328	2	Q6ZNM3_HUMAN	Q6ZNM3 homo sapien
25	18	60.0	332	2	Q6Z4E6_ORYSA	Q6Z4E6 oryza sativ
26	18	60.0	350	2	Q94589_9CILI	Q94589 leishadion b
27	18	60.0	355	1	RHOW_DROME	R20350 drosophila
28	18	60.0	355	2	Q540V7_DROME	Q540V7 drosophila
29	18	60.0	367	1	TISD_MOUSE	T23949 mus musculu
30	18	60.0	404	2	Q86SB6_DROVT	Q86SB6 drosophila
31	18	60.0	417	2	Q6Z4E5_ORYSA	Q6Z4E5 oryza sativ

32	18	60.0	445	2	Q8UUL4_ORYSA	Q8UUL4 oryzae lat
33	18	60.0	480	2	Q7SHC6_NEUTR	Q7SHC6 neurospora
34	18	60.0	480	2	Q758Y3_ASHGO	Q758Y3 ashya goss
35	18	60.0	492	1	TISD_HUMAN	P47974 homo sapien
36	18	60.0	494	2	Q53TB4_HUMAN	Q53TB4 homo sapien
37	18	60.0	526	2	Q6PD21_MOUSE	Q6PD21 mus musculu
38	18	60.0	544	2	Q5B195_DROME	Q5B195 drosophila
39	18	60.0	545	2	Q91TR2_TUHY1	Q91TR2 tupaiid her
40	18	60.0	551	2	Q5Z8N9_ORYSA	Q5Z8N9 oryza sativ
41	18	60.0	551	2	Q8SB30_ORYSA	Q8SB30 oryza sativ
42	18	60.0	552	2	Q5AP87_CANAL	Q5AP87 candida alb
43	18	60.0	560	2	Q80U53_HPBVO	Q80U53 hepatitis b
44	18	60.0	581	2	Q6N1S1_CORDI	Q6N1S1 corynebacte
45	18	60.0	592	2	Q51VZ0_MAGGR	Q51VZ0 magnaporthe
46	18	60.0	606	2	Q6Z4E7_ORYSA	Q6Z4E7 oryza sativ
47	18	60.0	691	2	Q4RVV9_TERNG	Q4RVV9 tetraodon n
48	18	60.0	705	2	Q7S136_NEUTR	Q7S136 neurospora
49	18	60.0	711	2	Q5RFRW2_BRAKE	Q5RFRW2 brachydanio
50	18	60.0	722	2	Q5AP84_CANAL	Q5AP84 candida alb
51	18	60.0	743	2	Q09517_HPBVO	Q09517 hepatitis b
52	18	60.0	769	2	Q80GX5_HPBVO	Q80GX5 hepatitis b
53	18	60.0	787	2	Q09511_HPBVO	Q09511 hepatitis b
54	18	60.0	795	2	Q8V1H4_HPBVO	Q8V1H4 hepatitis b
55	18	60.0	799	2	Q4T7H8_TERNG	Q4T7H8 tetraodon n
56	18	60.0	801	2	Q09504_HPBVO	Q09504 hepatitis b
57	18	60.0	821	2	Q09505_HPBVO	Q09505 hepatitis b
58	18	60.0	823	2	Q090M3_HPBVO	Q090M3 hepatitis b
59	18	60.0	836	2	Q8V1I3_HPBVO	Q8V1I3 hepatitis b
60	18	60.0	837	2	Q50UJ2_HPBVO	Q50UJ2 hepatitis b
61	18	60.0	843	1	DPOL_HPBVO	P17394 hepatitis b
62	18	60.0	843	1	DPOL_HPBVO	P17395 hepatitis b
63	18	60.0	843	2	Q09509_HPBVO	Q09509 hepatitis b
64	18	60.0	843	2	Q39877_HPBVO	Q39877 hepatitis b
65	18	60.0	843	2	Q39882_HPBVO	Q39882 hepatitis b
66	18	60.0	843	2	Q91527_HPBVO	Q91527 hepatitis b
67	18	60.0	843	2	Q5DW06_HPBVO	Q5DW06 hepatitis b
68	18	60.0	843	2	Q5DW09_HPBVO	Q5DW09 hepatitis b
69	18	60.0	843	2	Q5DW13_HPBVO	Q5DW13 hepatitis b
70	18	60.0	843	2	Q500T5_HPBVO	Q500T5 hepatitis b
71	18	60.0	843	2	Q67937_HPBVO	Q67937 hepatitis b
72	18	60.0	843	2	Q68RNL_HPBVO	Q68RNL hepatitis b
73	18	60.0	843	2	Q68RP0_HPBVO	Q68RP0 hepatitis b
74	18	60.0	843	2	Q68RP3_HPBVO	Q68RP3 hepatitis b
75	18	60.0	843	2	Q68RP6_HPBVO	Q68RP6 hepatitis b
76	18	60.0	843	2	Q68RQ9_HPBVO	Q68RQ9 hepatitis b
77	18	60.0	843	2	Q68RR3_HPBVO	Q68RR3 hepatitis b
78	18	60.0	843	2	Q762B4_HPBVO	Q762B4 hepatitis b
79	18	60.0	843	2	Q765V6_HPBVO	Q765V6 hepatitis b
80	18	60.0	843	2	Q76B37_HPBVO	Q76B37 hepatitis b
81	18	60.0	843	2	Q7TDP5_HPBVO	Q7TDP5 hepatitis b
82	18	60.0	843	2	Q7TDP8_HPBVO	Q7TDP8 hepatitis b
83	18	60.0	843	2	Q7TDP1_HPBVO	Q7TDP1 hepatitis b
84	18	60.0	843	2	Q7TDP1_HPBVO	Q7TDP1 hepatitis b
85	18	60.0	843	2	Q7TDP7_HPBVO	Q7TDP7 hepatitis b
86	18	60.0	843	2	Q7TDP8_HPBVO	Q7TDP8 hepatitis b
87	18	60.0	843	2	Q7TDP8_HPBVO	Q7TDP8 hepatitis b
88	18	60.0	843	2	Q7TDP8_HPBVO	Q7TDP8 hepatitis b
89	18	60.0	843	2	Q7TDP8_HPBVO	Q7TDP8 hepatitis b
90	18	60.0	843	2	Q7TDP8_HPBVO	Q7TDP8 hepatitis b
91	18	60.0	843	2	Q7TDP8_HPBVO	Q7TDP8 hepatitis b
92	18	60.0	843	2	Q805G6_HPBVO	Q805G6 hepatitis b
93	18	60.0	843	2	Q80GVR_HPBVO	Q80GVR hepatitis b
94	18	60.0	843	2	Q80H00_HPBVO	Q80H00 hepatitis b
95	18	60.0	843	2	Q80H20_HPBVO	Q80H20 hepatitis b
96	18	60.0	843	2	Q80H48_HPBVO	Q80H48 hepatitis b
97	18	60.0	843	2	Q80U60_HPBVO	Q80U60 hepatitis b
98	18	60.0	843	2	Q80J72_HPBVO	Q80J72 hepatitis b
99	18	60.0	843	2	Q80J78_HPBVO	Q80J78 hepatitis b
100	18	60.0	843	2	Q80J80_HPBVO	Q80J80 hepatitis b
101	18	60.0	843	2	Q80J83_HPBVO	Q80J83 hepatitis b
102	18	60.0	843	2	Q80M08_HPBVO	Q80M08 hepatitis b
103	18	60.0	843	2	Q81116_HPBVO	Q81116 hepatitis b
104	18	60.0	843	2	Q81120_HPBVO	Q81120 hepatitis b

105	18	60.0	843	2	Q81124	HBVVO	Q81124	hepatitis b	178	17	56.7	120	2	Q49859	MYCBE	Q49859	mycobacteri
106	18	60.0	843	2	Q8A240	HBVVO	Q8A240	hepatitis b	179	17	56.7	123	2	Q81A74	CABEL	Q81A74	caenorhabdi
107	18	60.0	843	2	Q8OSD1	HBVVO	Q8OSD1	hepatitis b	180	17	56.7	124	2	Q56R10	PENMO	Q56R10	penaeus mon
108	18	60.0	843	2	Q8VOM9	HBVVO	Q8VOM9	hepatitis b	181	17	56.7	126	2	Q56R14	GIALA	Q56R14	giardia lam
109	18	60.0	843	2	Q8VONI	HBVVO	Q8VONI	hepatitis b	182	17	56.7	126	2	Q9B119	GIALA	Q9B119	giardia lam
110	18	60.0	843	2	Q91F27	HBVVO	Q91F27	hepatitis b	183	17	56.7	126	2	Q9B1M3	GIALA	Q9B1M3	giardia lam
111	18	60.0	843	2	Q9QAB8	HBVVO	Q9QAB8	hepatitis b	184	17	56.7	127	2	Q5RQD0	ANOGA	Q5RQD0	anopheles g
112	18	60.0	843	2	Q9QAC4	HBVVO	Q9QAC4	hepatitis b	185	17	56.7	127	2	Q9B1K7	GIALA	Q9B1K7	giardia lam
113	18	60.0	843	2	Q9QAD4	HBVVO	Q9QAD4	hepatitis b	186	17	56.7	128	2	Q9B1K6	GIALA	Q9B1K6	giardia lam
114	18	60.0	843	2	Q9QAE0	HBVVO	Q9QAE0	hepatitis b	187	17	56.7	128	2	Q9B1K9	GIALA	Q9B1K9	giardia lam
115	18	60.0	843	2	Q9QBE2	HBVVO	Q9QBE2	hepatitis b	188	17	56.7	128	2	Q74AJO	GEOBFL	Q74AJO	geobacter s
116	18	60.0	843	2	Q9QBE7	HBVVO	Q9QBE7	hepatitis b	189	17	56.7	129	2	Q5VJF5	RAT	Q5VJF5	ratius norv
117	18	60.0	843	2	Q9QBF1	HBVVO	Q9QBF1	hepatitis b	190	17	56.7	130	1	Q9BHS5	HUMAN	Q9BHS5	homo sapien
118	18	60.0	843	2	Q9QBF6	HBVVO	Q9QBF6	hepatitis b	191	17	56.7	130	1	Q9BHS5	MOUSE	Q9BHS5	mus musculu
119	18	60.0	843	2	Q9YL94	HBVVO	Q9YL94	hepatitis b	192	17	56.7	130	2	Q6WNTD	HUMAN	Q6WNTD	homo sapien
120	18	60.0	843	2	Q6RZ64	HBVVO	Q6RZ64	hepatitis b	193	17	56.7	130	2	Q6AVB9	ORYSA	Q6AVB9	oryza sativ
121	18	60.0	843	2	Q7THS7	HBVVO	Q7THS7	hepatitis b	194	17	56.7	131	2	Q5TOR4	HUMAN	Q5TOR4	homo sapien
122	18	60.0	843	2	Q80U87	HBVVO	Q80U87	hepatitis b	195	17	56.7	132	2	Q8PKB7	XANAC	Q8PKB7	xanthomonas
123	18	60.0	843	2	Q5UAY3	HBVVO	Q5UAY3	hepatitis b	196	17	56.7	135	1	TVB1	HUMAN	TVB1	homo sapien
124	18	60.0	843	2	Q67925	HBVVO	Q67925	hepatitis b	197	17	56.7	139	2	Q5S508	PYRHO	Q5S508	pyrococcus
125	18	60.0	843	2	Q68RS2	HBVVO	Q68RS2	hepatitis b	198	17	56.7	140	2	Q6U7V1	CRYPE	Q6U7V1	crinipellis
126	18	60.0	843	2	Q6JWV7	HBVVO	Q6JWV7	hepatitis b	199	17	56.7	140	2	Q5BHK1	HUMAN	Q5BHK1	homo sapien
127	18	60.0	843	2	Q9PX62	HBVVO	Q9PX62	hepatitis b	200	17	56.7	140	2	Q6BDC0	HUMAN	Q6BDC0	homo sapien
128	18	60.0	843	2	Q9QAD0	HBVVO	Q9QAD0	hepatitis b	201	17	56.7	140	2	Q6F369	ORYSA	Q6F369	oryza sativ
129	18	60.0	843	2	Q9QAM9	HBVVO	Q9QAM9	hepatitis b	202	17	56.7	145	2	Q8VWK7	GLAMI	Q8VWK7	antirrhinum
130	18	60.0	843	2	Q9QMM9	HBVVO	Q9QMM9	hepatitis b	203	17	56.7	145	2	Q8W019	GLAMI	Q8W019	antirrhinum
131	18	60.0	843	2	Q5OUJ5	HBVVO	Q5OUJ5	hepatitis b	204	17	56.7	145	2	Q8W020	GLAMI	Q8W020	antirrhinum
132	18	60.0	843	2	Q4KRE9	HBVVO	Q4KRE9	hepatitis b	205	17	56.7	145	2	Q8W021	GLAMI	Q8W021	antirrhinum
133	18	60.0	909	2	Q52EM8	MAGGR	Q52EM8	magnaporthe	206	17	56.7	145	2	Q8W022	ANTMA	Q8W022	antirrhinum
134	18	60.0	984	2	Q4SMG0	TEYNG	Q4SMG0	tetradodon n	207	17	56.7	145	2	Q9A851	CANCR	Q9A851	caulobacter
135	18	60.0	1135	2	Q4PBA7	USTMA	Q4PBA7	usellago ma	208	17	56.7	145	2	Q9PDB1	BRAXE	Q9PDB1	brexydanto
136	18	60.0	1262	2	Q4RL33	TEYNG	Q4RL33	tetradodon n	209	17	56.7	146	2	Q9PDB1	BRAXE	Q9PDB1	brexydanto
137	18	60.0	1286	2	Q9VRA9	DROME	Q9VRA9	drosoophila	210	17	56.7	148	2	Q4UPP4	XANCP	Q4UPP4	xanthomonas
138	18	60.0	1348	2	Q4QUD0	LEIMA	Q4QUD0	leishmania	211	17	56.7	148	2	Q8P460	XANCP	Q8P460	xanthomonas
139	18	60.0	1349	2	Q675Y9	GUJOC	Q675Y9	oikopleura	212	17	56.7	150	2	Q7PT86	ANOGA	Q7PT86	anopheles g
140	18	60.0	1510	2	Q8PKK6	XANAC	Q8PKK6	xanthomonas	213	17	56.7	153	2	Q6WZR7	HUMAN	Q6WZR7	homo sapien
141	18	60.0	1743	2	Q9XXK5	CABEL	Q9XXK5	caenorhabdi	214	17	56.7	156	2	Q69S06	ORYSA	Q69S06	oryza sativ
142	18	60.0	2543	2	P90649	PARBR	P90649	paramecium	215	17	56.7	157	2	Q4LXB6	9BURK	Q4LXB6	burholderi
143	18	60.0	4261	2	Q81FP4	PLAF7	Q81FP4	plasmodium	216	17	56.7	158	2	Q9B1J9	GIALA	Q9B1J9	giardia lam
144	18	60.0	26	2	Q93940	PODAN	Q93940	podospora a	217	17	56.7	158	2	Q6Z4C9	ORYSA	Q6Z4C9	oryza sativ
145	17	56.7	26	2	Q7URM6	BOVIN	Q7URM6	bos taurus	218	17	56.7	161	2	Q9B1J8	GIALA	Q9B1J8	giardia lam
146	17	56.7	37	2	Q7URB3	RHOBA	Q7URB3	rhopodirell	219	17	56.7	161	2	Q4RZB3	TEYNG	Q4RZB3	tetradodon n
147	17	56.7	49	2	Q4MRU3	ASPFU	Q4MRU3	aspergillus	220	17	56.7	162	2	Q91GT6	ADEP3	Q91GT6	porcine ade
148	17	56.7	59	2	Q7SZA3	NEUCR	Q7SZA3	neurospora	221	17	56.7	164	2	Q77790	BOVIN	Q77790	bos taurus
149	17	56.7	62	2	Q84TW7	ORYSA	Q84TW7	oryza sativ	222	17	56.7	166	2	Q4W038	9BURK	Q4W038	burholderi
150	17	56.7	67	2	Q7QWX1	GIALA	Q7QWX1	giardia lam	223	17	56.7	168	2	Q9ZKFS	RHIME	Q9ZKFS	rhizobium va
151	17	56.7	83	2	Q7R273	GIALA	Q7R273	giardia lam	224	17	56.7	174	1	ITRY	BAUVA	ITRY	baubinia va
152	17	56.7	84	2	Q8GMP9	ARATH	Q8GMP9	arabidopsis	225	17	56.7	175	2	Q9S1Z1	SHEBP	Q9S1Z1	ovis aries
153	17	56.7	84	2	Q8L9N7	ARATH	Q8L9N7	arabidopsis	226	17	56.7	178	2	Q7NTE5	CHRYVO	Q7NTE5	chromopacte
154	17	56.7	85	2	Q5WM46	LEBPL	Q5WM46	legionella	227	17	56.7	179	2	Q61TV1	CABER	Q61TV1	caenorhabdi
155	17	56.7	85	2	Q5X4Q7	LEBPA	Q5X4Q7	legionella	228	17	56.7	181	2	Q6Z6C6	ORYSA	Q6Z6C6	oryza sativ
156	17	56.7	85	2	Q5ZUZ4	LEBGP	Q5ZUZ4	legionella	229	17	56.7	182	1	YAP2	CABEL	YAP2	caenorhabdi
157	17	56.7	86	2	Q8DMN7	VIBVU	Q8DMN7	vibrio vuln	230	17	56.7	184	2	Q61ZB0	CABER	Q61ZB0	caenorhabdi
158	17	56.7	86	2	Q7MGN7	VIBVU	Q7MGN7	vibrio vuln	231	17	56.7	184	2	Q7Z2D2	DESEVR	Q7Z2D2	desulffovibr
159	17	56.7	87	2	Q5UTM7	ORYSA	Q5UTM7	oryza sativ	232	17	56.7	185	2	Q800G0	TEYNG	Q800G0	tetradodon n
160	17	56.7	88	2	Q74DU1	GEOBL	Q74DU1	geobacter s	233	17	56.7	188	2	Q5S6G2	HUMAN	Q5S6G2	homo sapien
161	17	56.7	88	2	Q90ZP3	XENLA	Q90ZP3	xenopus lae	234	17	56.7	188	2	Q90XY9	SCYCA	Q90XY9	scyliothinu
162	17	56.7	89	2	Q49652	MYCBE	Q49652	mycobacteri	235	17	56.7	194	2	Q86NU8	MOUSE	Q86NU8	mus musculu
163	17	56.7	95	2	Q7YTR5	CABEL	Q7YTR5	caenorhabdi	236	17	56.7	195	1	Q5USW6	MOUSE	Q5USW6	mus musculu
164	17	56.7	95	2	P94628	CITPR	P94628	citrobacter	237	17	56.7	196	2	Q5M988	XENLA	Q5M988	xenopus lae
165	17	56.7	96	2	Q7YTR6	CABEL	Q7YTR6	caenorhabdi	238	17	56.7	196	2	Q4SMD6	TEYNG	Q4SMD6	tetradodon n
166	17	56.7	96	2	Q4SDL9	TEYNG	Q4SDL9	tetradodon n	239	17	56.7	196	2	Q7R5C3	GIALA	Q7R5C3	giardia lam
167	17	56.7	99	1	NICI	HUMAN	Q9UG19	homo sapien	240	17	56.7	200	2	Q7R5C3	GIALA	Q7R5C3	giardia lam
168	17	56.7	99	1	Q61AD7	HUMAN	Q61AD7	homo sapien	241	17	56.7	202	2	P95044	MYCTU	P95044	mycobacteri
169	17	56.7	99	2	Q5YA4U	ABEOR	Q5YA4U	agelena ori	242	17	56.7	204	2	Q4LHS3	9BURK	Q4LHS3	burholderi
170	17	56.7	101	2	Q48435	BSPBP	Q48435	bacterioph	243	17	56.7	205	2	Q5TOW7	ANOGA	Q5TOW7	anopheles g
171	17	56.7	102	2	Q5A3U4	CANAL	Q5A3U4	candida alb	244	17	56.7	205	2	Q94R93	TEYNG	Q94R93	tetradodon n
172	17	56.7	102	2	Q4S1V5	TEYNG	Q4S1V5	tetradodon n	245	17	56.7	209	2	Q94R28	TEYTH	Q94R28	tetradodon n
173	17	56.7	108	2	Q5DAH3	SCHUA	Q5DAH3	schistosoma	246	17	56.7	210	2	Q4P148	USTMA	Q4P148	usellago ma
174	17	56.7	108	2	Q5DCP1	SCHUA	Q5DCP1	schistosoma	247	17	56.7	210	2	Q4R8B9	DEIRA	Q4R8B9	tetradodon n
175	17	56.7	108	2	Q7SHC3	ORYSA	Q7SHC3	oryza sativ	248	17	56.7	212	2	Q9RRS8	DEIRA	Q9RRS8	delnoccoc
176	17	56.7	108	2	Q53CR5	9GAMA	Q53CR5	macaca fusc	249	17	56.7	213	2	Q800F5	TEYNG	Q800F5	tetradodon n
177	17	56.7	109	2	Q7U1F7	MYCBO	Q7U1F7	mycobacteri	250	17	56.7	215	2	Q9HC44	HUMAN	Q9HC44	homo sapien

251	17	56.7	216	2	Q4T5X5_TETNG	Q4T5X5 tetradon n	324	17	56.7	292	1	CITG_ECO57	P58161 escherichia
252	17	56.7	217	2	Q8BV88_MOISE	Q8BV88 mus musculus	325	17	56.7	292	1	CITG_ECOL6	Q8BF04 escherichia
253	17	56.7	218	2	Q95WY8_IKOSC	Q95WY8 ixodes scap	326	17	56.7	292	1	CITG_ECOL1	P77231 escherichia
254	17	56.7	219	2	Q4N1W0_THRPA	Q4N1W0 thelteria p	327	17	56.7	292	2	Q7UDF6_SHIFL	Q7UDF6 shigella fl
255	17	56.7	220	2	Q4PMW5_IKOSC	Q4PMW5 ixodes scap	328	17	56.7	293	2	Q6NLK5_DROME	Q6NLK5 drosophila
256	17	56.7	220	2	Q4PMK0_IKOSC	Q4PMK0 ixodes scap	329	17	56.7	294	2	Q6ZON2_CABBR	Q6ZON2 caenorhabdi
257	17	56.7	221	1	KR103_HUMAN	P60369 homo sapien	330	17	56.7	294	2	Q615G2_CABBR	Q615G2 caenorhabdi
258	17	56.7	223	1	Q7MXL2_ALCEU	Q7MXL2 alcaligenes	331	17	56.7	294	2	Q4RPT5_TETNG	Q4RPT5 tetradon n
259	17	56.7	224	1	MYOG_HUMAN	P15173 homo sapien	332	17	56.7	296	2	Q5PEG9_SALPA	Q5PEG9 salmonella
260	17	56.7	224	1	MYOG_PIG	P49812 sus scrofa	333	17	56.7	296	2	Q73TV0_MYCPA	Q73TV0 mycobacteri
261	17	56.7	224	2	Q53XW6_HUMAN	Q53XW6 homo sapien	334	17	56.7	297	2	Q4P724_USTMA	Q4P724 ussillago ma
262	17	56.7	224	2	Q564X4_CABBL	Q564X4 caenorhabdi	335	17	56.7	297	2	Q4FX78_LEIMA	Q4FX78 leishmania
263	17	56.7	224	2	Q5EB03_PIG	Q5EB03 sus scrofa	336	17	56.7	297	2	Q8DA53_AGRY5	Q8DA53 agrobacteri
264	17	56.7	224	2	Q7YS81_BOVIN	Q7YS81 bos taurus	337	17	56.7	299	2	Q7Q3P3_ANOGA	Q7Q3P3 anopheles g
265	17	56.7	224	2	Q5EC23_PIG	Q5EC23 sus scrofa	338	17	56.7	299	2	Q4VSE5_DROME	Q4VSE5 drosophila
266	17	56.7	225	2	Q6BUR5_DEBHA	Q6BUR5 debaryomyce	339	17	56.7	300	2	Q4PEP7_USTMA	Q4PEP7 ussillago ma
267	17	56.7	225	2	Q6YXSG_ORYSA	Q6YXSG oryza sativ	340	17	56.7	301	2	Q5RQK9_ANOGA	Q5RQK9 anopheles g
268	17	56.7	226	2	Q5NTG7_AZOSE	Q5NTG7 azoarcus sp	341	17	56.7	301	2	Q8BN70_MOUSE	Q8BN70 mus musculu
269	17	56.7	227	2	Q5N970_ORYSA	Q5N970 oryza sativ	342	17	56.7	302	1	HXC1A_BRARE	Q98B19 brachydantio
270	17	56.7	227	2	Q51IQ3_VIRIU	Q51IQ3 banana viru	343	17	56.7	302	2	Q7XH78_ORYSA	Q7XH78 oryza sativ
271	17	56.7	230	1	UL16_HGMVA	P16757 human cytom	344	17	56.7	302	2	Q8WSM4_ORYSA	Q8WSM4 oryza sativ
272	17	56.7	231	2	Q5TMZ9_ANOGA	Q5TMZ9 anopheles g	345	17	56.7	302	2	Q5FOJ6_GLTOX	Q5FOJ6 gluconobact
273	17	56.7	235	2	Q8UTX8_XENIA	Q8UTX8 xenopus lae	346	17	56.7	304	2	Q5TMU6_ANOGA	Q5TMU6 anopheles g
274	17	56.7	236	2	Q5DZAS_VIBF1	Q5DZAS vibrio flic	347	17	56.7	304	2	Q5TKI6_SALCH	Q5TKI6 salmonella
275	17	56.7	236	2	Q8UXU7_XENIA	Q8UXU7 xenopus lae	348	17	56.7	305	2	YGBT_ECOL1	Q61WAS phobacter
276	17	56.7	237	2	Q41N40_GIBZE	Q41N40 gibberella	349	17	56.7	305	2	Q4XQ66_PLACH	Q4XQ66 plasmodium
277	17	56.7	237	2	Q92XP6_RHIME	Q92XP6 rhizobium m	350	17	56.7	305	2	Q8Z467_SALTI	Q8Z467 salmonella
278	17	56.7	238	1	RNRH_RHINI	P08056 rhizopus ni	351	17	56.7	306	2	Q61174_CABBR	Q61174 caenorhabdi
279	17	56.7	238	2	Q81128_CABBL	Q81128 caenorhabdi	352	17	56.7	306	2	Q5TKI6_SALCH	Q5TKI6 salmonella
280	17	56.7	240	2	Q5TSH9_ANOGA	Q5TSH9 anopheles g	353	17	56.7	306	2	Q61WAS_PHOFR	Q61WAS phobacter
281	17	56.7	240	2	Q41ZX3_9BURK	Q41ZX3 burkholderi	354	17	56.7	306	2	Q8ZMP2_SALTY	Q8ZMP2 salmoneila
282	17	56.7	242	2	Q5TPG2_ANOGA	Q5TPG2 anopheles g	355	17	56.7	307	2	Q74DC4_GEOSL	Q74DC4 geobacter s
283	17	56.7	242	2	Q8EGH3_SHEON	Q8EGH3 shewanella	356	17	56.7	307	2	Q8X7W3_ECO57	Q8X7W3 escherichia
284	17	56.7	243	1	TRP1_XENIA	P19799 xenopus lae	357	17	56.7	307	2	Q8VGO3_MOUSE	Q8VGO3 mus musculu
285	17	56.7	243	2	Q5M959_XENTR	Q5M959 xenopus tto	358	17	56.7	309	2	Q17234_CABBL	Q17234 caenorhabdi
286	17	56.7	243	2	Q5M976_XENTR	Q5M976 xenopus tto	359	17	56.7	311	2	Q9V684_DROME	Q9V684 drosophila
287	17	56.7	243	2	Q7S206_XENIA	Q7S206 xenopus lae	360	17	56.7	314	2	Q7RSR0_PLAYO	Q7RSR0 plasmodium
288	17	56.7	243	2	Q8AV83_BRARE	Q8AV83 brachydantio	361	17	56.7	315	2	Q5SRK3_HUMAN	Q5SRK3 homo sapien
289	17	56.7	244	1	KLK6_HUMAN	Q92876 homo sapien	362	17	56.7	315	2	Q97601_RABIT	Q97601 coryctolagus
290	17	56.7	244	2	Q6H301_HUMAN	Q6H301 homo sapien	363	17	56.7	317	2	Q7NVD5_CHRVO	Q7NVD5 chromopacte
291	17	56.7	244	2	Q9YXZ6_LEIMA	Q9YXZ6 leishmania	364	17	56.7	319	2	Q5SGCB3_HUMAN	Q5SGCB3 homo sapien
292	17	56.7	247	2	Q966B6_CABBL	Q966B6 caenorhabdi	365	17	56.7	319	2	Q4VSL2_DROME	Q4VSL2 drosophila
293	17	56.7	247	2	Q561Z7_BRARE	Q561Z7 brachydantio	366	17	56.7	319	2	Q4S7N3_TETNG	Q4S7N3 tetradon n
294	17	56.7	248	1	CUTC_PORGI	Q7mbd6 porphyromon	367	17	56.7	320	2	Q8BNV3_CORGL	Q8BNV3 corynebacte
295	17	56.7	248	2	Q61HM9_CABBR	Q61HM9 caenorhabdi	368	17	56.7	326	2	Q4SHX8_TETNG	Q4SHX8 tetradon n
296	17	56.7	248	2	Q91LZX3_ARATH	Q91LZX3 arabidopsis	369	17	56.7	330	2	Q6WPO7_9ETKA	Q6WPO7 tetradicho
297	17	56.7	250	2	Q6HX41_BACAN	Q6HX41 bacillus an	370	17	56.7	333	2	Q6CKQ7_XENIA	Q6CKQ7 xenopus lae
298	17	56.7	251	1	KR10A_HUMAN	P60014 homo sapien	371	17	56.7	334	2	Q4LHM6_9BURK	Q4LHM6 burkholderi
299	17	56.7	252	2	Q8RXT8_ARATH	Q8RXT8 arabidopsis	372	17	56.7	347	2	Q26995_TOXGO	Q26995 toxoplasma
300	17	56.7	253	2	Q5CAQ0_CHICK	Q5CAQ0 gallus gall	373	17	56.7	347	2	Q8Z9C0_STREB	Q8Z9C0 streptomyce
301	17	56.7	255	1	KR102_HUMAN	P60368 homo sapien	374	17	56.7	348	2	Q9S213_STREB	Q9S213 streptomyce
302	17	56.7	259	2	Q52T86_VIBCH	Q52T86 vibrio chol	375	17	56.7	350	1	LVM2_ARATH	Q4S1B9 tetradon n
303	17	56.7	259	2	Q4LY98_9BURK	Q4LY98 burkholderi	376	17	56.7	351	2	Q4S1B9_TETNG	Q4S1B9 tetradon n
304	17	56.7	260	2	Q7Y150_ORYSA	Q7Y150 oryza sativ	377	17	56.7	352	2	Q5GVU1_XANOR	Q5GVU1 xanthomonas
305	17	56.7	266	2	Q73QF7_TREDE	Q73QF7 treponema d	378	17	56.7	352	2	Q9A610_CAVCR	Q9A610 caulobacter
306	17	56.7	267	2	Q9ABR0_CAVCR	Q9ABR0 caulobacter	379	17	56.7	358	1	MRAY_NOCFA	Q5YYU2 nocardiacta fa
307	17	56.7	271	1	KR105_HUMAN	P60370 homo sapien	380	17	56.7	361	2	Q6HH39_BACIK	Q6HH39 bacillus th
308	17	56.7	271	1	MURI_MYCBO	P63536 mycobacteri	381	17	56.7	361	2	Q73SR6_BACIK	Q73SR6 bacillus ce
309	17	56.7	271	1	MURI_MYCTU	P63535 mycobacteri	382	17	56.7	361	2	Q81NM8_BACAN	Q81NM8 bacillus an
310	17	56.7	272	1	MURI_MYCLR	P46705 mycobacteri	383	17	56.7	363	2	Q639Q7_BACIZ	Q639Q7 bacillus ce
311	17	56.7	275	2	Q73X85_MYCPA	Q73X85 mycobacteri	384	17	56.7	363	2	Q5D8F2_SCHVA	Q5D8F2 schistosoma
312	17	56.7	276	2	Q5MIT5_ABDAL	Q5MIT5 aedes albop	385	17	56.7	363	2	Q41WH8_AZCVI	Q41WH8 azobacter
313	17	56.7	277	2	Q7MPJ3_VIBVY	Q7MPJ3 vibrio vuln	386	17	56.7	363	2	Q4UXI9_CORJK	Q4UXI9 corynebacte
314	17	56.7	278	2	Q7S1H4_ORYSA	Q7S1H4 oryza sativ	387	17	56.7	365	1	KR106_HUMAN	Q63971 homo sapien
315	17	56.7	280	2	Q6R8Z3_CIOIN	Q6R8Z3 cioma intes	388	17	56.7	365	2	Q8FDG2_ECOL6	Q8FDG2 escherichia
316	17	56.7	282	1	KR101_HUMAN	P60331 homo sapien	389	17	56.7	365	2	Q83712_TREPA	Q83712 treponema p
317	17	56.7	283	2	Q4RKF2_TETNG	Q4RKF2 tetradon n	390	17	56.7	366	2	Q5S562_NPLAS	Q5S562 leucantia se
318	17	56.7	283	2	Q5AKZ6_DICDI	Q5AKZ6 dicystosell	391	17	56.7	367	2	Q8NSA6_HUMAN	Q8NSA6 homo sapien
319	17	56.7	283	2	Q4SUT1_TETNG	Q4SUT1 tetradon n	392	17	56.7	367	2	Q6AVJ3_ORYSA	Q6AVJ3 oryza sativ
320	17	56.7	284	2	Q91295_RANCA	Q91295 rana catesb	393	17	56.7	370	1	KR107_HUMAN	P60409 homo sapien
321	17	56.7	286	2	Q4S0U6_TETNG	Q4S0U6 tetradon n	394	17	56.7	371	2	Q6Z6N5_CABBR	Q6Z6N5 caenorhabdi
322	17	56.7	290	2	Q7Z729_HUMAN	Q7Z729 homo sapien	395	17	56.7	371	2	Q4RQ43_TETNG	Q4RQ43 tetradon n
323	17	56.7	290	2	Q6XPZ1_SCYCA	Q6XPZ1 scyllorhinu	396	17	56.7	372	2	Q63PT9_BURPS	Q63PT9 burkholderi

397	17	56.7	373	2	Q80YN6_RAT	Q80YN6_rattus novy	470	17	56.7	461	2	Q81ZK8_HUMAN	Q81ZK8_homo sapien
398	17	56.7	375	2	Q7NGJ0_GLOVI	Q7NGJ0_gloeobacter	471	17	56.7	461	2	Q9BQ80_HUMAN	Q9BQ80_homo sapien
399	17	56.7	376	2	Q97G11_CIOAB	Q97G11_clostridium	472	17	56.7	462	2	Q8WQ6_CABEL	Q8WQ6_caenorhabdi
400	17	56.7	377	2	Q96V44_TRIRE	Q96V44_trichoderma	473	17	56.7	462	2	Q4NQPS_9BELT	Q4NQPS_aeareomyxob
401	17	56.7	378	2	Q9KJL9_TRISI	Q9KJL9_trichoderma	474	17	56.7	467	2	Q584K9_9TRYP	Q584K9_9tryp
402	17	56.7	381	2	Q81G20_CABEL	Q81G20_caenorhabdi	475	17	56.7	467	2	Q17233_BOOMO	Q17233_bombux mori
403	17	56.7	381	2	Q4NZ90_9DELT	Q4NZ90_aeareomyxob	476	17	56.7	467	2	Q4R3M6_MACFA	Q4R3M6_macaca fasc
404	17	56.7	381	2	Q4LQ26_9BORK	Q4LQ26_burkholderi	477	17	56.7	470	2	Q7QWMS_GIALA	Q7QWMS_giardia lam
405	17	56.7	382	2	Q60C97_METCA	Q60C97_methylcoccc	478	17	56.7	471	2	VS1L_TRYBB	P26326_9trypnosoma
406	17	56.7	383	2	Q4SY38_TETNG	Q4SY38_tetradodon n	479	17	56.7	471	2	VSMA_TRYBR	P20946_9trypnosoma
407	17	56.7	384	2	Q7SEPF_9ASHG	Q7SEPF_9ashya gos	480	17	56.7	472	2	Q4FYNN_LEIMA	Q4FYNN_leishmania
408	17	56.7	385	2	Q7SEPF_9ASHG	Q7SEPF_9ashya gos	481	17	56.7	472	2	Q5XHA0_XENTR	Q5XHA0_xenopus tro
409	17	56.7	388	2	Q95W43_ENCTN	Q95W43_nocardi	482	17	56.7	473	2	Q7TTO6_XENTL	Q7TTO6_xenopus lae
410	17	56.7	388	2	Q5YNC4_NOCFA	Q5YNC4_nocardi	483	17	56.7	475	2	Q60UP6_CABER	Q60UP6_caenorhabdi
411	17	56.7	391	2	Q74BN0_GEOSL	Q74BN0_geobacter s	484	17	56.7	475	2	Q61UP8_CABER	Q61UP8_caenorhabdi
412	17	56.7	393	2	Q6AL17_DESPS	Q6AL17_desulfocale	485	17	56.7	476	2	Q9EL54_HPVAT	Q9EL54_hepatitis b
413	17	56.7	394	2	Q72EB4_DESVH	Q72EB4_desulfovibr	486	17	56.7	477	2	Q519Q4_ENTHI	Q519Q4_entamoeba h
414	17	56.7	395	2	Q95V70_TETTH	Q95V70_tetrahymena	487	17	56.7	478	2	Q4W8S1_HUMAN	Q4W8S1_homo sapien
415	17	56.7	396	1	IP6K3_MOUSE	Q8BQd2_mus musculu	488	17	56.7	478	2	Q9W487_DROME	Q9W487_drosophila
416	17	56.7	396	1	ISCS_RUMFL	Q54055_ruminococcu	489	17	56.7	478	2	Q8S999_PHANT	Q8S999_phaeocylus a
417	17	56.7	396	2	Q4LGT2_9BURK	Q4LGT2_burkholderi	490	17	56.7	485	2	Q58559_9XANT	Q58559_xanthomonas
418	17	56.7	397	2	Q95V71_TETTH	Q95V71_tetrahymena	491	17	56.7	486	2	Q6WRD8_BDEBA	Q6WRD8_dbelliovibri
419	17	56.7	398	1	NIRS2_ANAVA	Q44482_anabaena va	492	17	56.7	487	2	Q4S723_TETNG	Q4S723_tetradodon n
420	17	56.7	400	1	NIRS_ANAAZ	Q43884_anabaena az	493	17	56.7	487	2	MDM4_MOUSE	Q35618_mus musculu
421	17	56.7	400	1	NIRS_ANASP	P12623_anabaena sp	494	17	56.7	489	1	Q9CYG1_MOUSE	Q9CYG1_mus muscu
422	17	56.7	400	2	Q09256_SYNP8	Q09256_synchococc	495	17	56.7	489	2	MDM4_HUMAN	Q15151_homo sapien
423	17	56.7	400	2	Q896F7_CLOTE	Q896F7_clostridium	496	17	56.7	490	2	Q6GS18_HUMAN	Q6GS18_homo sapien
424	17	56.7	401	1	KR104_HUMAN	P60372_homo sapien	497	17	56.7	490	2	Q5XINI_RAT	Q5XINI_rattus novy
425	17	56.7	402	2	Q62FW8_BUMA	Q62FW8_burkholderi	498	17	56.7	490	2	Q99L86_MOUSE	Q99L86_mus musculu
426	17	56.7	404	2	Q88LX5_PSPRK	Q88LX5_pseudomonas	499	17	56.7	492	1	E1BL_ADE07	P03345_human adeno
427	17	56.7	405	2	Q5VKR9_SACER	Q5VKR9_baccharopol	500	17	56.7	492	2	Q5EY82_ADE07	Q5EY82_human adeno
428	17	56.7	407	2	Q511AI_MAGAR	Q511AI_magnaporthe	501	17	56.7	492	2	Q6PBT9_BRABE	Q6PBT9_brachydanio
429	17	56.7	407	2	Q8B9B2_NPVRO	Q8B9B2_rachidiplusia	502	17	56.7	494	2	Q571X1_TRYBR	Q571X1_9trypnosoma
430	17	56.7	408	1	VLEN_NPVAC	P24647_autographia	503	17	56.7	494	2	Q74AB6_GEOSL	Q74AB6_geobacter s
431	17	56.7	410	1	IP6K3_HUMAN	Q95PC2_homo sapien	504	17	56.7	498	2	Q4PBP5_USTMA	Q4PBP5_ustilago ma
432	17	56.7	410	2	Q5TER7_HUMAN	Q5TER7_homo sapien	505	17	56.7	500	2	Q8N4S1_HUMAN	Q8N4S1_homo sapien
433	17	56.7	410	2	Q5TRQ4_HUMAN	Q5TRQ4_homo sapien	506	17	56.7	502	2	Q6UD12_9HERP	Q6UD12_gibberella
434	17	56.7	410	2	Q5RC09_PONPY	Q5RC09_pongo pygma	507	17	56.7	506	2	Q51XK4_MAGAR	Q51XK4_magnaporthe
435	17	56.7	410	2	Q7NZY8_CHRVO	Q7NZY8_chromobacte	508	17	56.7	510	2	Q4IGF6_GIBZE	Q4IGF6_gibberella
436	17	56.7	411	2	Q6CGF0_YARLI	Q6CGF0_yarrowia ll	509	17	56.7	514	2	Q8H812_ORYZA	Q8H812_oryza sativ
437	17	56.7	411	2	Q820T4_ENTFA	Q820T4_enterococcu	510	17	56.7	515	2	Q8UUY9_BRABE	Q8UUY9_brachydanio
438	17	56.7	412	2	Q8VB76_BRIME	Q8VB76_brucella me	511	17	56.7	515	2	Q5THX6_HUMAN	Q5THX6_homo sapien
439	17	56.7	413	2	Q4QD02_LEIMA	Q4QD02_leishmania	512	17	56.7	517	2	Q4W8S3_HUMAN	Q4W8S3_homo sapien
440	17	56.7	413	2	Q4RH65_TETNG	Q4RH65_tetradodon n	513	17	56.7	518	2	Q94718_PARTE	Q94718_parametium
441	17	56.7	415	2	Q8GE34_HEIMO	Q8GE34_heliobacilli	514	17	56.7	521	2	Q80XU0_MOUSE	Q80XU0_mus musculu
442	17	56.7	419	2	Q4QDN4_LEIMA	Q4QDN4_leishmania	515	17	56.7	522	1	B028_CABEL	P41396_mus musculu
443	17	56.7	421	2	Q95V69_TETTH	Q95V69_tetrahymena	516	17	56.7	524	2	Q6PFI6_HUMAN	Q6PFI6_homo sapien
444	17	56.7	423	2	Q8OQW7_METMA	Q8OQW7_methanosarc	517	17	56.7	524	2	Q60MX6_CABER	Q60MX6_caenorhabdi
445	17	56.7	423	2	Q9U697_TETTH	Q9U697_tetrahymena	518	17	56.7	524	2	Q41507_GIBZE	Q41507_gibberella
446	17	56.7	423	2	Q7XK63_ORYSA	Q7XK63_oryza sativ	519	17	56.7	525	2	Q41815_GIBZE	Q41815_gibberella
447	17	56.7	431	2	Q21650_CABEL	Q21650_caenorhabdi	520	17	56.7	527	2	Q41815_GIBZE	Q41815_gibberella
448	17	56.7	432	2	Q4Q476_LEIMA	Q4Q476_leishmania	521	17	56.7	530	1	TRAPE_MOUSE	P70196_mus musculu
449	17	56.7	435	2	Q616G8_CABER	Q616G8_caenorhabdi	522	17	56.7	530	2	Q7U6V5_SYNPX	Q7U6V5_synchococc
450	17	56.7	436	2	Q27197_TETTH	Q27197_tetrahymena	523	17	56.7	532	2	Q6P9M0_MOUSE	Q6P9M0_mus musculu
451	17	56.7	437	2	Q5UIK5_ORYSA	Q5UIK5_oryza sativ	524	17	56.7	532	2	Q9SK50_ARYTH	Q9SK50_arabidopsis
452	17	56.7	439	2	Q8AXV2_FUGRU	Q8AXV2_fugu rubrip	525	17	56.7	539	2	Q7R020_GIALA	Q7R020_giardia lam
453	17	56.7	439	2	Q94827_TETTH	Q94827_tetrahymena	526	17	56.7	543	2	Q9UVL7_PNECA	Q9UVL7_pneumocysti
454	17	56.7	439	2	Q8JID1_BRABE	Q8JID1_brachydanio	527	17	56.7	543	2	Q4P677_USTMA	Q4P677_ustilago ma
455	17	56.7	440	2	Q8N1H1_HUMAN	Q8N1H1_homo sapien	528	17	56.7	546	2	Q68EW3_XENTL	Q68EW3_xenopus lae
456	17	56.7	444	2	Q4QBT5_LEIMA	Q4QBT5_leishmania	529	17	56.7	547	1	C78A1_MAIZB	P48420_zea mays (m
457	17	56.7	445	2	Q9NC96_9TRYP	Q9NC96_9trypnosoma	530	17	56.7	547	2	Q41295_GIBZE	Q41295_gibberella
458	17	56.7	445	2	Q57XL9_9TRYP	Q57XL9_9trypnosoma	531	17	56.7	547	2	Q4T338_TETNG	Q4T338_tetradodon n
459	17	56.7	447	2	Q96FD7_HUMAN	Q96FD7_homo sapien	532	17	56.7	549	2	Q968Y9_HUMAN	Q968Y9_homo sapien
460	17	56.7	447	2	Q8PLV4_XANAC	Q8PLV4_xanthomonas	533	17	56.7	553	2	Q68FS8_PARTE	Q68FS8_parametium
461	17	56.7	449	2	Q53RR6_HUMAN	Q53RR6_homo sapien	534	17	56.7	555	1	Q5QYAI_PIG	Q5QYAI_sus scrofa
462	17	56.7	453	2	Q51P64_MAGAR	Q51P64_magnaporthe	535	17	56.7	555	2	SY14_HUMAN	P34424_caenorhabdi
463	17	56.7	454	2	Q8C1S8_MOUSE	Q8C1S8_mus musculu	536	17	56.7	555	1	YL13_CABEL	Q5THX7_homo sapien
464	17	56.7	456	1	RINI_RAT	P29315_rattus novy	537	17	56.7	556	2	Q6DUN2_XENTL	Q6DUN2_xenopus lae
465	17	56.7	456	2	Q924P4_MOUSE	Q924P4_mus musculu	538	17	56.7	556	2	Q5MRY3_LEGPT	Q5MRY3_legionella
466	17	56.7	456	2	Q91V17_MOUSE	Q91V17_mus musculu	539	17	56.7	567	2	Q91XCI_ARYTH	Q91XCI_arabidopsis
467	17	56.7	460	1	RINI_HUMAN	P13489_homo sapien	541	17	56.7	574	2	Q5THX8_HUMAN	Q5THX8_homo sapien
468	17	56.7	460	2	Q871X0_NEUCR	Q871X0_neutrospora	542	17	56.7	577	1	ITB6_CAVPO	P18563_cavia porce

543	17	56.7	582	2	062JU0_BURMA	062JU0	burkholderi
544	17	56.7	585	2	091F49_HPBVO	091f49	hepatitis b
545	17	56.7	596	2	05VTL7_HUMAN	05vlt7	homo sapien
546	17	56.7	596	2	08T6A8_CAEEL	08t6a8	caenorhabditi
547	17	56.7	596	2	061TR2_RAT	061tr2	rattus norv
548	17	56.7	601	2	08LM57_ORYSA	08lm57	oryza sativ
549	17	56.7	602	2	04P9J9_USITWA	04p9j9	ustiliago ma
550	17	56.7	603	2	06ZUR2_ORYSA	06zur2	oryza sativ
551	17	56.7	603	2	04NNZ1_9DELT	04nnz1	anaeromyxob
552	17	56.7	604	2	09NKG5_HUMAN	09nkg5	homo sapien
553	17	56.7	608	2	07YI66_CRYPV	07yi66	cryptospori
554	17	56.7	609	2	052GAI_MAGGR	052gai	magnaporthe
555	17	56.7	609	2	05CPJ1_CRYPV	05cpj1	cryptospori
556	17	56.7	612	2	063UV6_BURPS	063uv6	burkholderi
557	17	56.7	612	2	04SBD5_TETNG	04sbd5	tetradodon n
558	17	56.7	613	2	04N940_THERP	04n940	theileria p
559	17	56.7	615	2	09LOP9_ARATH	09lop9	arabidopsis
560	17	56.7	615	2	07SSR7_NEUCR	07ssr7	neurospora
561	17	56.7	629	2	07ZTW7_BRARE	07ztw7	brachydanio
562	17	56.7	631	2	0485Z3_ARATH	0485z3	arabidopsis
563	17	56.7	638	2	09LOP8_ARATH	09lop8	arabidopsis
564	17	56.7	640	2	09N5X0_CAEEL	09n5x0	caenorhabditi
565	17	56.7	650	2	05EBE8_BRARE	05ebe8	brachydanio
566	17	56.7	653	2	04KMG1_HUMAN	04kmg1	homo sapien
567	17	56.7	654	2	08QZ20_MOUSE	08qz20	mus musculu
568	17	56.7	655	2	09H0T6_HUMAN	09h0t6	homo sapien
569	17	56.7	656	2	05ARH5_DICDI	05arh5	dicrocoeli
570	17	56.7	656	2	05U2T7_RAT	05u2t7	rattus norv
571	17	56.7	658	2	018215_CAEEL	018215	caenorhabditi
572	17	56.7	659	2	05ZCG0_ORYSA	05zcg0	oryza sativ
573	17	56.7	663	2	052D83_MAGGR	052d83	magnaporthe
574	17	56.7	667	1	TS11.GIALA	003185	giardia lam
575	17	56.7	670	2	0627F3_CAEER	0627f3	caenorhabditi
576	17	56.7	670	2	0774Z3_BACBR	0774z3	bacrocera
577	17	56.7	674	2	04PLN4_USITWA	04pln4	ustiliago ma
578	17	56.7	674	2	052KX8_XENLA	052kx8	xenopus lae
579	17	56.7	676	2	051W16_9REOV	051w16	avian ortho
580	17	56.7	677	1	WHITE_BACCU	005360	lucilia cup
581	17	56.7	679	2	081S30_BACCU	081s30	bactrocera
582	17	56.7	681	2	072Z29_BRARE	072z29	brachydanio
583	17	56.7	682	2	07YOK2_BOVIN	07yok2	bos taurus
584	17	56.7	684	2	081498_CUPSA	081498	cuplemnius
585	17	56.7	684	2	07ZVU9_BRARE	07zv9	brachydanio
586	17	56.7	687	1	WHITE_DROME	P10090	drosophila
587	17	56.7	687	2	094960_DROSI	094960	drosophila
588	17	56.7	691	2	04QB67_LEIMA	04qb67	leishmania
589	17	56.7	695	2	07LSY4_HUMAN	07lsy4	homo sapien
590	17	56.7	695	2	05BD07_BOVIN	05bd07	bos taurus
591	17	56.7	705	2	04TVD1_TETNG	04tvd1	tetradodon n
592	17	56.7	707	2	06CAV7_YARLI	06cav7	yarrowia li
593	17	56.7	707	2	08BW61_MOUSE	08bw61	m mus muscu
594	17	56.7	712	2	04FX57_LEIMA	04fx57	leishmania
595	17	56.7	715	2	05RDL9_PONPY	05rdl9	pongo pygma
596	17	56.7	719	2	09U019_GIALA	09u019	giardia lam
597	17	56.7	719	2	09U021_GIALA	09u021	giardia lam
598	17	56.7	725	2	07YOK1_BOVIN	07yok1	bos taurus
599	17	56.7	725	2	08BZM8_MOUSE	08bzm8	mus musculu
600	17	56.7	726	2	08BRF2_HUMAN	08brf2	homo sapien
601	17	56.7	737	2	08BW83_MOUSE	08bw83	m mus muscu
602	17	56.7	746	2	08S6L6_ORYSA	08s6l6	oryza sativ
603	17	56.7	747	2	08CAV1_MOUSE	08cav1	mus musculu
604	17	56.7	747	2	06NWS7_BRARE	06nws7	brachydanio
605	17	56.7	748	2	08N264_HUMAN	08n264	homo sapien
606	17	56.7	748	2	09CW64_MOUSE	09cw64	mus musculu
607	17	56.7	749	2	04LT50_9BURK	04lt50	burkholderi
608	17	56.7	751	2	08LRK4_ARATH	08lrk4	arabidopsis
609	17	56.7	751	2	04MO64_9BURK	04mo64	burkholderi
610	17	56.7	769	2	04P3X8_USITWA	04p3x8	ustiliago ma
611	17	56.7	783	2	05WT26_LEGGL	05wt26	legionella
612	17	56.7	783	2	05X1B3_LEGGL	05x1b3	legionella
613	17	56.7	783	2	05ZR04_LEGPH	05zr04	legionella
614	17	56.7	788	1	ITB6_HUMAN	P18564	homo sapien
615	17	56.7	794	2	09U110_LEIMA	09u110	leishmania
616	17	56.7	796	2	05TV91_HUMAN	05tv91	homo sapien
617	17	56.7	807	2	05TV06_9TRYP	05tv06	trypanosoma
618	17	56.7	817	2	06T6B3_CHICK	06t6b3	gallus gall
619	17	56.7	813	2	05DVV9_AERRO	05dvv9	aeromonas s
620	17	56.7	814	2	05DVW0_AERRO	05dvw0	aeromonas s
621	17	56.7	820	2	080YV4_MOUSE	080yv4	mus musculu
622	17	56.7	823	2	04SUN7_TETNG	04sun7	tetradodon n
623	17	56.7	823	2	04S9S7_TETNG	04s9s7	tetradodon n
624	17	56.7	833	2	06J288_ACAÇA	06j288	acanthamoeb
625	17	56.7	836	2	080J66_HPBVO	080j66	hepatitis b
626	17	56.7	843	1	DPOL_HPBVI	P17353	hepatitis b
627	17	56.7	843	2	068RR7_HPBVO	068rr7	hepatitis b
628	17	56.7	843	2	068H36_HPBVO	068h36	hepatitis b
629	17	56.7	843	2	09QMI9_HPBVO	09qmi9	hepatitis b
630	17	56.7	843	2	09QMJ3_HPBVO	09qmj3	hepatitis b
631	17	56.7	843	2	09YLP1_HPBVO	09y1p1	hepatitis b
632	17	56.7	843	2	003766_HPBVO	003766	hepatitis b
633	17	56.7	844	2	05H1Y6_XANOR	05h1y6	xanthomonas
634	17	56.7	849	2	04PBQ2_USITWA	04pbq2	ustiliago ma
635	17	56.7	849	2	05YQX9_NOCFA	05yqx9	nocardia fa
636	17	56.7	853	2	08VXP5_MYCTU	08vxp5	mycobacteri
637	17	56.7	853	2	07UXG2_MYCBO	07uxg2	mycobacteri
638	17	56.7	853	2	053674_MYCTU	053674	mycobacteri
639	17	56.7	854	2	05ZBC9_ORYSA	05zbc9	oryza sativ
640	17	56.7	855	2	06GVN7_MYCAV	06gvn7	mycobacteri
641	17	56.7	855	2	07JTL6_MYCPA	07jtl6	mycobacteri
642	17	56.7	855	2	06FBZ5_ACIAD	06fbz5	acinetobact
643	17	56.7	864	2	04SY29_TETNG	04sy29	tetradodon n
644	17	56.7	870	2	06PCD4_HUMAN	06pcd4	homo sapien
645	17	56.7	873	2	04S9L8_TETNG	04s9l8	tetradodon n
646	17	56.7	879	2	09XZ14_DROME	09xz14	drosophila
647	17	56.7	893	1	RUSC1_MOUSE	08bq66	mus musculu
648	17	56.7	894	2	08NC14_HUMAN	08nc14	homo sapien
649	17	56.7	894	2	05RER7_PONPY	05rer7	pongo pygma
650	17	56.7	900	2	05XKF8_CRYNE	05xkf8	cryptococcu
651	17	56.7	902	2	05KMH5_CRYNE	05kmh5	cryptococcu
652	17	56.7	902	1	RUSC1_HUMAN	09bvn2	homo sapien
653	17	56.7	902	2	05TV22_HUMAN	05tv22	homo sapien
654	17	56.7	903	2	055XB7_CRYNE	055xb7	cryptococcu
655	17	56.7	903	2	05KML6_CRYNE	05kml6	cryptococcu
656	17	56.7	904	1	TLR3_HUMAN	015455	homo sapien
657	17	56.7	904	2	05Q4W0_HUMAN	05q4w0	homo sapien
658	17	56.7	904	2	04VAL2_HUMAN	04val2	homo sapien
659	17	56.7	909	2	06MP91_BDEBA	06mp91	bdellovibri
660	17	56.7	912	2	055N08_CRYNE	055n08	cryptococcu
661	17	56.7	912	2	05KCA3_CRYNE	05kca3	cryptococcu
662	17	56.7	921	2	04UGL3_THEAN	04ugl3	theileria a
663	17	56.7	929	2	05TVH8_9TRYP	05tvh8	trypanosoma
664	17	56.7	934	1	CO6_HUMAN	P13671	homo sapien
665	17	56.7	934	1	CO6_PANTR	P61135	pan troglod
666	17	56.7	934	1	CO6_PONPY	P61135	pongo pygma
667	17	56.7	938	2	07YTA2_CRYPV	07yta2	cryptospori
668	17	56.7	938	2	05CLY0_CRYHO	05cly0	cryptospori
669	17	56.7	938	2	04K9C6_PSEPS	04k9c6	pseudomonas
670	17	56.7	950	2	08RZX0_ORYSA	08rx0	oryza sativ
671	17	56.7	969	2	07RE04_GIALA	07re04	giardia lam
672	17	56.7	975	2	08NCC9_HUMAN	08ncc9	homo sapien
673	17	56.7	983	2	05HBA4_HUMAN	05hba4	homo sapien
674	17	56.7	993	2	04QOG9_LEIMA	04qog9	leishmania
675	17	56.7	1002	2	04SQO0_TETNG	04sqo0	tetradodon n
676	17	56.7	1002	2	095WA4_ENCIN	095wa4	encephalito
677	17	56.7	1003	2	08IMQ3_DROME	08imq3	drosophila
678	17	56.7	1011	2	04S0D2_TETNG	04s0d2	tetradodon n
679	17	56.7	1021	2	04TB22_TETNG	04tb22	tetradodon n
680	17	56.7	1023	2	08SB11_ORYSA	08sb11	oryza sativ
681	17	56.7	1023	2	07XGB0_ORYSA	07xgb0	oryza sativ
682	17	56.7	1028	2	076836_CAEEL	076836	caenorhabditi
683	17	56.7	1031	2	06JNA7_LYCES	06jna7	lycopersico
684	17	56.7	1049	2	04Y187_PLACH	04y187	plasmodium
685	17	56.7	1059	2	08CHP0_MOUSE	08chp0	mus musculu
686	17	56.7	1065	1	LRTG2_HUMAN	094598	homo sapien
687	17	56.7	1080	1	HDC_DROME	09nrm8	drosophila
688	17	56.7	1089	2	08T3A0_CIOIN	08t3a0	ciona intes



835	16	53.3	69	2	Q9NCT8_9CAEN	Q9NCT8 conus ebrae	908	16	53.3	102	2	Q6PA15_MOUSE	Q6PA15 mus musculu
836	16	53.3	69	2	Q9NCT9_9CAEN	Q9NCT9 conus ebrae	909	16	53.3	103	2	Q07106_YEAST	Q07106 saccharomyc
837	16	53.3	69	2	Q9NCT0_9CAEN	Q9NCT0 conus ebrae	910	16	53.3	103	2	Q7XOH3_ORYZA	Q7XOH3 oryza sativ
838	16	53.3	69	2	Q9RIH3_XENNE	Q9RIH3 xenorhabdus	911	16	53.3	104	2	Q7Z576_HUMAN	Q7Z576 homo sapien
839	16	53.3	70	2	Q4LBEC2_HUMAN	Q4LBEC2 homo sapien	912	16	53.3	104	2	Q9NCR5_9CUTU	Q9NCR5 dendrocid
840	16	53.3	70	2	Q8XTW3_RALSO	Q8XTW3 ralestoma s	913	16	53.3	104	2	Q9UT34_NEIMA	Q9UT34 neiberia m
841	16	53.3	71	1	CXO6A_CONCT	CXO6A conus catus	914	16	53.3	104	2	Q5YFC9_9VIRU	Q5YFC9 groupere ixi
842	16	53.3	71	2	Q8VB36_MSSV	Q8VB36 white spot	915	16	53.3	104	2	Q5GAE2_9VIRU	Q5GAE2 leishmania
843	16	53.3	72	2	Q6YSG0_ORYSA	Q6YSG0 oryza sativ	916	16	53.3	105	2	Q4OEY4_LEIMA	Q4OEY4 leishmania
844	16	53.3	73	2	Q9P9X6_XYLPA	Q9P9X6 xyella faa	917	16	53.3	105	2	Q64QV5_BACR	Q64QV5 bacteroides
845	16	53.3	74	2	Q4RC41_TETNG	Q4RC41 tetraodon n	918	16	53.3	105	2	Q5W7P6_BRARE	Q5W7P6 brachydanio
846	16	53.3	75	2	Q8CIN8_MOUSE	Q8CIN8 mus musculu	919	16	53.3	106	2	Q4M0G8_BURK	Q4M0G8 burkholderi
847	16	53.3	76	2	Q4KD95_PSEBS	Q4KD95 pseudomonas	920	16	53.3	106	2	Q4V2B3_BURKA	Q4V2B3 burkholderi
848	16	53.3	77	2	Q5OMJ7_ORYSA	Q5OMJ7 oryza sativ	921	16	53.3	107	2	Q8MXU4_CABEL	Q8MXU4 caenorhabdi
849	16	53.3	77	2	Q8H387_ORYSA	Q8H387 oryza sativ	922	16	53.3	107	2	Q68800_PSEBE	Q68800 pseudomonas
850	16	53.3	77	2	Q7P334_PUSNV	Q7P334 fusobacteri	923	16	53.3	108	1	PRV4_MERMR	P02621 merlangius
851	16	53.3	78	2	Q804C5_CHICK	Q804C5 gallus gall	924	16	53.3	108	2	Q7SB84_NEUR	Q7SB84 neurospora
852	16	53.3	78	2	Q4SWR4_TETNG	Q4SWR4 tetraodon n	925	16	53.3	108	2	Q8WPO3_9TRYP	Q8WPO3 trypanosoma
853	16	53.3	79	1	CMC17_EIMAC	C21959 elmeria ace	926	16	53.3	109	2	Q46346_9CUCU	Q46346 dendrocid
854	16	53.3	79	2	Q8MTB6_SCHUA	Q8MTB6 schistosoma	927	16	53.3	109	2	Q9YHM6_GINCI	Q9YHM6 ginglymasto
855	16	53.3	79	2	Q87AZ1_XYLPA	Q87AZ1 xyella faa	928	16	53.3	109	2	Q90YL0_GADMO	Q90YL0 gadus morhu
856	16	53.3	79	2	Q9PGA0_XYLPA	Q9PGA0 xyella faa	929	16	53.3	110	2	Q7PRG2_ANOGA	Q7PRG2 anopheles g
857	16	53.3	79	2	Q4U747_XYLEB	Q4U747 prochlorodis	930	16	53.3	110	2	Q06095_MYCIE	Q06095 mycobacteri
858	16	53.3	80	2	Q4T1S0_TETNG	Q4T1S0 tetraodon n	931	16	53.3	110	2	Q8G611_BIFLO	Q8G611 bifidobacte
859	16	53.3	81	2	Q57C12_BRUBA	Q57C12 bruceella ab	932	16	53.3	111	2	Q54XP7_DICDI	Q54XP7 dictyostell
860	16	53.3	81	2	Q5LC78_BACRN	Q5LC78 bacteroides	933	16	53.3	111	2	Q65MLJ3_ORYSA	Q65MLJ3 oryza sativ
861	16	53.3	81	2	Q64T67_BACPR	Q64T67 bacteroides	934	16	53.3	111	2	Q5SML3_ORYSA	Q5SML3 oryza sativ
862	16	53.3	81	2	Q6VB63_HMYR	Q6VB63 human herpe	935	16	53.3	111	2	Q8GB95_MYCAY	Q8GB95 mycobacteri
863	16	53.3	82	2	Q7Y5B3_BPR69	Q7Y5B3 bacteriopho	936	16	53.3	112	2	Q4SWR5_TETNG	Q4SWR5 tetraodon n
864	16	53.3	82	2	Q9XPU0_9CHLO	Q9XPU0 chlamydomo	937	16	53.3	112	2	Q9U747_TENNO	Q9U747 tenerio mo
865	16	53.3	82	2	Q5KWM2_GEOXA	Q5KWM2 geobacillius	938	16	53.3	112	2	Q4SVY7_TETNG	Q4SVY7 tenerio mo
866	16	53.3	82	2	Q82PU6_STRAW	Q82PU6 streptomyce	939	16	53.3	112	2	Q6MP17_YAVES	Q6MP17 yarrowia l
867	16	53.3	83	1	VIREB3_AGRYS	P08061 agrobacteri	940	16	53.3	113	2	Q6C958_YARLI	Q6C958 yarrowia l
868	16	53.3	83	2	Q687A6_9ZYCO	Q687A6 parasitella	941	16	53.3	113	2	Q59115_PYRHO	Q59115 pyrococcus
869	16	53.3	83	2	Q9REB4_9RHIZ	Q9REB4 agrobacteri	942	16	53.3	114	2	Q8EPJ7_XANAC	Q8EPJ7 xanthomonas
870	16	53.3	83	2	Q5P6Q2_AZOSE	Q5P6Q2 azoarcus sp	943	16	53.3	114	2	Q8AXG3_GINCI	Q8AXG3 ginglymsto
871	16	53.3	83	2	Q728G9_DESYH	Q728G9 desulfovibr	944	16	53.3	114	2	Q4T4H7_TETNG	Q4T4H7 tetraodon n
872	16	53.3	83	2	P97880_RAT	P97880 rattus norv	945	16	53.3	115	1	CVPS_PIMHY	Q81G02 caenorhabdi
873	16	53.3	84	2	Q6G1U1_9DIPY	Q6G1U1 culicoides	946	16	53.3	115	1	Q6CAK3_ARATH	P20254 pseudochis
874	16	53.3	84	2	Q6GZD0_PRG3V	Q6GZD0 frog virus	947	16	53.3	115	2	Q8CBZ5_MORBE	P20257 pseudochis
875	16	53.3	84	2	Q6YH96_9VIRU	Q6YH96 ambystoma t	948	16	53.3	115	2	Q8CBZ5_MORBE	Q8CBZ5 mus musculu
876	16	53.3	85	2	Q5P617_AZOSE	Q5P617 azoarcus sp	949	16	53.3	115	2	Q7TFD7_RHOM6	Q7TFD7 rhesus cyto
877	16	53.3	85	2	Q6LHP2_PHOBR	Q6LHP2 photobacter	950	16	53.3	115	2	Q8JGK4_GINCI	Q8JGK4 ginglymsto
878	16	53.3	86	2	Q87KG3_YIBPA	Q87KG3 vibrio para	951	16	53.3	117	2	Q81G02_CABEL	Q81G02 caenorhabdi
879	16	53.3	86	2	Q4RAZ0_TETNG	Q4RAZ0 burkholderi	952	16	53.3	118	1	PA20_PSEAU	P20254 pseudochis
880	16	53.3	87	2	Q4LH80_9BURK	Q4LH80 burkholderi	953	16	53.3	118	1	PA21B_PSEAU	P20254 pseudochis
881	16	53.3	87	2	Q6L220_PICRO	Q6L220 picophilus	954	16	53.3	118	1	PA22_ACAN	P81227 acanthophis
882	16	53.3	88	2	Q6M1A4_RHISN	Q6M1A4 rhizobium s	955	16	53.3	118	1	PA29_PSEAU	P20253 pseudochis
883	16	53.3	89	2	Q9BGV8_MACRA	Q9BGV8 macaca fasc	956	16	53.3	118	1	PA2C_PSEAU	P20256 pseudochis
884	16	53.3	90	2	Q5J178_ORYSA	Q5J178 oryza sativ	957	16	53.3	118	1	PA2D_PSEAU	P20257 pseudochis
885	16	53.3	90	2	Q5WYU7_LEGFL	Q5WYU7 legionella	958	16	53.3	118	1	PA2E_PSEAU	P20257 pseudochis
886	16	53.3	90	2	Q5X776_LEGPA	Q5X776 legionella	959	16	53.3	118	2	Q5UT00_ANOGA	Q5UT00 anopheles g
887	16	53.3	90	2	Q5ZKX8_LEGPH	Q5ZKX8 legionella	960	16	53.3	118	2	Q9NCU7_9CUCU	Q9NCU7 dendrocid
888	16	53.3	91	2	Q4XC49_PLACH	Q4XC49 plasmodium	961	16	53.3	119	1	PA21B_ACAN	P81229 acanthophis
889	16	53.3	91	2	Q400G5_IPONI	Q400G5 ipomoea nil	962	16	53.3	119	1	PA21B_NOTSC	P00607 notschis sc
890	16	53.3	91	2	Q8X3G1_ECOS7	Q8X3G1 escherichia	963	16	53.3	119	2	Q6UXS6_HUMAN	Q6UXS6 homo sapien
891	16	53.3	91	2	Q42468_IPONI	Q42468 ipomoea nil	964	16	53.3	119	2	Q4UTW5_XANCP	Q4UTW5 xanthomonas
892	16	53.3	92	2	Q98198_MCVI	Q98198 molluscum c	965	16	53.3	119	2	Q97KS4_CLOAB	Q97KS4 clostridium
893	16	53.3	92	2	Q69373_9ALPH	Q69373 cercopithec	966	16	53.3	119	2	Q8JGAE6_GINCI	Q8JGAE6 ginglymsto
894	16	53.3	93	2	Q6OLX3_CABER	Q6OLX3 caenorhabdi	967	16	53.3	120	2	Q5G6W8_9CUCU	Q5G6W8 microdera d
895	16	53.3	96	2	Q98EAB_RHIO	Q98EAB rhizobium l	968	16	53.3	120	2	Q9B1K3_GIALA	Q9B1K3 agardia lam
896	16	53.3	97	2	Q81XK0_DROME	Q81XK0 drosophila	969	16	53.3	120	2	Q67423_AOUAE	Q67423 aguilae aeo
897	16	53.3	98	2	Q91V05_MOUSE	Q91V05 mus musculu	970	16	53.3	121	2	Q6S0P8_DROPS	Q6S0P8 drosophila
898	16	53.3	98	2	Q6PA14_MOUSE	Q6PA14 mus musculu	971	16	53.3	121	2	Q9NCR1_9CUCU	Q9NCR1 synchococc
899	16	53.3	99	2	Q4S6H4_TETNG	Q4S6H4 tetraodon n	972	16	53.3	121	2	Q5M207_SYNP6	Q5M207 burkholderi
900	16	53.3	100	2	Q6EPZ4_ORYSA	Q6EPZ4 oryza sativ	973	16	53.3	121	2	Q63104_BURPS	Q63104 burkholderi
901	16	53.3	100	2	Q8HYV9_BRUME	Q8HYV9 bruceella me	974	16	53.3	122	2	Q9NCR4_9CUCU	Q9NCR4 dendrocid
902	16	53.3	101	2	Q9SRS1_DROME	Q9SRS1 drosophila	975	16	53.3	122	2	Q6RH05_BACHK	Q6RH05 bacillus th
903	16	53.3	101	2	Q8C3R3_MOUSE	Q8C3R3 mus musculu	976	16	53.3	122	2	Q73D11_BACCI	Q73D11 bacillus ce
904	16	53.3	102	2	Q94F05_HORVU	Q94F05 hordeum vul	977	16	53.3	122	2	Q81HH9_BACCR	Q81HH9 bacillus ce
905	16	53.3	102	2	Q6S737_PICAB	Q6S737 picea abies	978	16	53.3	122	2	Q81UQ2_BACAN	Q81UQ2 bacillus an
906	16	53.3	102	2	Q9D702_MOUSE	Q9D702 mus musculu	979	16	53.3	122	2		
907	16	53.3	102	2			980	16	53.3	122	2		



981 16 53.3 122 2 063f19\_BACCC 063f19 bacillus ce  
982 16 53.3 122 2 05UVZ1\_GADBN 05UVZ1 human adeno  
983 16 53.3 122 2 071923\_ADB35 071923 human adeno  
984 16 53.3 122 2 08B8S0\_GADBN 08B8S0 human adeno  
985 16 53.3 123 1 IL21\_PERMA 080x92 peromyces  
986 16 53.3 123 2 07P38\_ANOGA 07P38 anopheles g  
987 16 53.3 123 2 08IG94\_GINCI 08IG94 ginglymsto  
988 16 53.3 124 2 08PV06\_METMA 08PV06 methanarc  
989 16 53.3 124 2 07R266\_GIALA 07R266 giardia lam  
990 16 53.3 124 2 016121\_TENMO 016121 tenebrio mo  
991 16 53.3 124 2 062710\_CAEBR 062710 caenorhabd  
992 16 53.3 124 2 09U744\_TENMO 09U744 tenebrio mo  
993 16 53.3 125 2 06DLX5\_TENMO 06DLX5 tenebrio mo  
994 16 53.3 125 2 08I4U7\_CABEU 08I4U7 caenorhabd  
995 16 53.3 125 2 038346\_BPLIH 038346 laetococcus  
996 16 53.3 125 2 04NR72\_9DELT 04NR72 anaeromyxob  
997 16 53.3 125 2 08JFJ4\_GINCI 08JFJ4 ginglymsto  
998 16 53.3 125 2 08JFJ7\_GINCI 08JFJ7 ginglymsto  
999 16 53.3 125 2 08JFJ8\_GINCI 08JFJ8 ginglymsto  
1000 16 53.3 125 2 08JFJ0\_GINCI 08JFJ0 ginglymsto

## ALIGNMENTS

RESULT 1  
Q9N669\_9CAEN PRELIMINARY; PRT; 69 AA.

ID Q9N669\_9CAEN PRELIMINARY; PRT; 69 AA.  
AC Q9N669;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
DE Four-loop conotoxin (Fragment).  
OS *Conus ebraeus*.  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxId=89425;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20416305; PubMed=10958845;  
RA Duda T.F., Jr., Palumbi S.R.;  
RT "Evolutionary diversification of multigene families: allelic selection  
of toxins in predatory cone snails.";  
RL Mol. Biol. Evol. 17:1286-1293(2000).  
RL [2]  
RN NUCLEOTIDE SEQUENCE.  
RP Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.  
RA EMBL; AF174281; AAF89945.1; -; mRNA.  
DR EMBL; AF174280; AAF89944.1; -; mRNA.  
DR GO; GO:0005576; C:extracellular region; IEA.  
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR004214; Conotoxin.  
DR Pfam; PF02950; Conotoxin; 1.  
FT NON\_TER 1  
SQ SEQUENCE 69 AA; 7662 MW; CEC826FDC920C4F1 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 69;  
Best Local Similarity 25.0%; Pred. No. 2e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 61 CSTATSTC 68

RESULT 2  
Q9NCS8\_9CAEN PRELIMINARY; PRT; 69 AA.  
ID Q9NCS8\_9CAEN PRELIMINARY;  
AC Q9NCS8;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Four-loop conotoxin (Fragment).  
OS *Conus ebraeus*.  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxId=89425;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20416305; PubMed=10958845;  
RA Duda T.F., Jr., Palumbi S.R.;  
RT "Evolutionary diversification of multigene families: allelic selection  
of toxins in predatory cone snails.";  
RL Mol. Biol. Evol. 17:1286-1293(2000).  
RL [2]  
RN NUCLEOTIDE SEQUENCE.  
RP Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.  
RA EMBL; AF174287; AAF89951.1; -; mRNA.  
DR GO; GO:0005576; C:extracellular region; IEA.  
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR004214; Conotoxin.  
DR Pfam; PF02950; Conotoxin; 1.  
FT NON\_TER 1  
SQ SEQUENCE 69 AA; 7674 MW; 83D526FDC934422A CRC64;

Query Match 60.0%; Score 18; DB 2; Length 69;  
Best Local Similarity 25.0%; Pred. No. 2e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
DB 61 CSTATSTC 68

RESULT 3  
Q9NCS9\_9CAEN PRELIMINARY; PRT; 69 AA.  
ID Q9NCS9\_9CAEN PRELIMINARY;  
AC Q9NCS9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Four-loop conotoxin (Fragment).  
OS *Conus ebraeus*.  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCBI\_TaxId=89425;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20416305; PubMed=10958845;  
RA Duda T.F., Jr., Palumbi S.R.;  
RT "Evolutionary diversification of multigene families: allelic selection  
of toxins in predatory cone snails.";  
RL Mol. Biol. Evol. 17:1286-1293(2000).  
RL [2]  
RN NUCLEOTIDE SEQUENCE.  
RP Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.  
RA EMBL; AF174286; AAF89950.1; -; mRNA.  
DR GO; GO:0005576; C:extracellular region; IEA.  
DR GO; GO:0008200; F:ion channel inhibitor activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR004214; Conotoxin.  
DR Pfam; PF02950; Conotoxin; 1.  
FT NON\_TER 1  
SQ SEQUENCE 69 AA; 7650 MW; DF3826FDC920C4F1 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 69;  
Best Local Similarity 25.0%; Pred. No. 2e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;



QY 4 CXXXXXXC 11  
|  
Db 61 CSTATSTC 68

## RESULT 4

Q9NCT2\_9CAEN PRELIMINARY; PRT; 69 AA.  
ID Q9NCT2\_9CAEN PRELIMINARY;  
AC Q9NCT2;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DE Four-loop conotoxin (Fragment).  
OS Conus ebraeus.  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
OC Neogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
OC NCB1\_TaxID=89425;  
OX NCB1\_TaxID=89425;  
RN NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20416305; PubMed=10958845;  
RA Duda T.F., Jr., Palumbi S.R.,  
RT "Evolutionary diversification of multigene families: allelic selection  
of toxins in predatory cone snails.";  
RL Mol. Biol. Evol. 17:1286-1293(2000).  
[2]

QY 4 CXXXXXXC 11  
|  
Db 61 CSTATSTC 68

## RESULT 5

Q9NCT3\_9CAEN PRELIMINARY; PRT; 69 AA.  
ID Q9NCT3\_9CAEN PRELIMINARY;  
AC Q9NCT3;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DE Four-loop conotoxin (Fragment).  
OS Conus ebraeus.  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
OC Neogastropoda; Conoidea; Conidae; Conus.  
OX NCB1\_TaxID=89425;  
RN NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20416305; PubMed=10958845;  
RA Duda T.F., Jr., Palumbi S.R.,  
RT "Evolutionary diversification of multigene families: allelic selection  
of toxins in predatory cone snails.";  
RL Mol. Biol. Evol. 17:1286-1293(2000).  
[2]

QY 4 CXXXXXXC 11  
|  
Db 61 CSTATSTC 68

## RESULT 6

Q9VD36\_DROME PRELIMINARY; PRT; 69 AA.  
ID Q9VD36\_DROME PRELIMINARY;  
AC Q9VD36;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE CG13858-PA.  
GN Name=CG13858; ORFName=CG13858;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCB1\_TaxID=7227;  
RN NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gockayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,  
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davernport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laske P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Martel B., McIntosh T.C., Mcleod M.P., McPherson D.,  
RA Merkulov G., Mikhlin N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitak R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
[2]

DR GO:0005576; C:extracellular region; IEA.  
DR GO:0008200; F:ion channel inhibitor activity; IEA.  
DR GO:0009405; P:pathogenesis; IEA.  
DR InterPro: IPR004214; Conotoxin.  
DR Pfam: PF02950; Conotoxin; 1.  
FT NON TER 1  
SQ SEQUENCE 69 AA; 7643 MW; D5699CFDC920CAFL CRC64;

Query Match 60.0%; Score 18; DB 2; Length 69;  
Best Local Similarity 25.0%; Pred. No. 2e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
|  
Db 61 CSTATSTC 68

Q9VD36\_DROME PRELIMINARY; PRT; 69 AA.  
ID Q9VD36\_DROME PRELIMINARY;  
AC Q9VD36;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE CG13858-PA.

GN Name=CG13858; ORFName=CG13858;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCB1\_TaxID=7227;  
RN NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gockayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,  
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davernport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laske P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Martel B., McIntosh T.C., Mcleod M.P., McPherson D.,  
RA Merkulov G., Mikhlin N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitak R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
[2]

```

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celnikier S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodegryn E.J.,
RA Svirskas R., Taber P.B., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.B., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RL melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celnikier S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomic perspective."
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,
RA Bettencourt B.R., Celnikier S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RL systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celnikier S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX Flybase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003738; AAF55967.1; -; Genomic DNA.
DR Ensembl; CG13858; Drosophila melanogaster.
DR Flybase; FBgn0040585; CG13858.
SQ SEQUENCE 69 AA; 6492 MW; EB077B26FD260BCC CRC64;

Query Match 60.0%; Score 18; DB 2; Length 69;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 9 CSASASC 16

RESULT 7
Q70FP9_ANOGA PRELIMINARY; PRT; 90 AA.
AC Q70FP9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP00000019154 (Fragment).
GN ORFNames=ENSANGG00000016655;
OS Anopheles gambiae str. PST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
NCBI_TaxID=180454;

```

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008846; EAA06382.2; -; Genomic DNA.
FT NON TER 90
SQ SEQUENCE 90 AA; 9710 MW; 72DB5798DF140ADF CRC64;

Query Match 60.0%; Score 18; DB 2; Length 90;
Best Local Similarity 25.0%; Pred. No. 2.2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 20 CSASASC 27

RESULT 8
Q9B111_ANOGA PRELIMINARY; PRT; 90 AA.
AC Q9B111;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE GSG1a protein precursor (Fragment).
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GASUA; TISSUE=Salivary glands;
RX MEDLINE=22057806; PubMed=12062411; DOI=10.1016/S0014-5793(02)02578-4;
RA Lantancotti A., Lombardo F., Santolamazza F., Veneri M.,
RA Castiglione T., Coluzzi M., Arca B.;
RT "Novel cDNAs encoding salivary proteins from the malaria vector
RT Anopheles gambiae."
RL FEBS Lett. 517:67-71(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GASUA; TISSUE=Salivary glands;
RA Arca B.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ297932; CAC35452.1; -; mRNA.
DR Ensembl; ENSANGG00000016655; Anopheles gambiae.
KW Signal.
FT SIGNAL 1 25 Potential;
FT NON TER 90
SQ SEQUENCE 90 AA; 9710 MW; 72DB5798DF140ADF CRC64;

Query Match 60.0%; Score 18; DB 2; Length 90;
Best Local Similarity 25.0%; Pred. No. 2.2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 20 CSASASC 27

RESULT 9
Q82854_9RETR PRELIMINARY; PRT; 114 AA.
ID Q82854_9RETR

```

AC Q82854; 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE TAR protein (Transactivating regulatory protein).  
 GN Name=tat;  
 OS Jembrana disease virus.  
 OC Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;  
 OC Lentiviruses; Bovine lentiviruses.  
 NC NCB1\_TaxID=36370;  
 RX MEDLINE=97201584; PubMed=9049370;  
 RX Chaddick B.J., Coelen R.J., Wilcox G.E., Samuels L.M., Kertayadnya G.;  
 "Nucleotide sequence analysis of Jembrana disease virus: a bovine  
 lentivirus associated with an acute disease syndrome.";  
 J. Gen. Virol. 76:1637-1650 (1995).  
 DR EMBL; U21603; AAA64395.1; -; Genomic RNA.  
 DR GO; GO:0042025; C:host cell nucleus; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR GO; GO:0006350; P:transcription; IEA.  
 DR InterPro: IPR001831; IV\_Tat.  
 DR Pfam: PF00539; Tat; 1.  
 KW activator; Nuclear protein; RNA-binding; Transcription;  
 KW Transcription regulation.  
 SQ SEQUENCE 114 AA; 12457 MW; 77A7C6B6FAF128D5A CRC64;

Query Match 60.0%; Score 18; DB 2; Length 114;  
 Best Local Similarity 25.0%; Pred. No. 2.5e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 96 CASSASC 103

RESULT 10  
 Q9EOL8 MOUSE PRELIMINARY; PRT; 161 AA.  
 ID Q9EOL8; 01-MAR-2001 (TrEMBLrel. 16, Created)  
 AC Q9EOL8;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE TIS1D deletion variant (Fragment).  
 GN Name=zfp3612; Synonyms=Brf2;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 NC NCB1\_TaxID=10090;  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BLKS/J;  
 RA Cho K., Hobson K., Greenhalgh D.G.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF205222; AAG45250.1; -; mRNA.  
 DR HSSP; P22893; IM90.  
 DR SMR; Q9EOL8; 1-49.  
 DR MGI; MGI:107945; zfp3612.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro: IPR000571; Znf\_CCCH.  
 DR Pfam; PF00642; zf-CCCH; 1.  
 DR SMART; SM00356; Znf\_C3H1; 1.  
 FT NON\_TER 1 1  
 FT SEQUENCE 161 AA; 17101 MW; BF09DF73709215EE CRC64;

Query Match 60.0%; Score 18; DB 2; Length 161;  
 Best Local Similarity 25.0%; Pred. No. 2.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 4 CXXXXXXC 11  
 DB 129 CASSASC 136

RESULT 11  
 Q7QPM4 GIALA PRELIMINARY; PRT; 163 AA.  
 ID Q7QPM4; 01-MAR-2004 (TrEMBLrel. 26, Created)  
 AC Q7QPM4;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE GIP 54 16521 16030.  
 OS Giardia lamblia ATCC 50803.  
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
 NC NCB1\_TaxID=184922;  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=WB C6;  
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,  
 RA Olsen G.J., Sogin M.L.;  
 RT "Draft sequence of the Giardia lamblia genome.";  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 preliminary data.  
 DR EMBL; AACB0100166; EAA36969.1; -; Genomic DNA.  
 SQ SEQUENCE 163 AA; 17478 MW; 3AF4A5F8EDA9A10C CRC64;

Query Match 60.0%; Score 18; DB 2; Length 163;  
 Best Local Similarity 25.0%; Pred. No. 2.8e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 24 CAATTAAC 31

RESULT 12  
 Q8LDT9 ARATH PRELIMINARY; PRT; 177 AA.  
 ID Q8LDT9; 01-OCT-2002 (TrEMBLrel. 22, Created)  
 AC Q8LDT9;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 NC NCB1\_TaxID=3702;  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22088475; PubMed=12093376;  
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,  
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
 RT "Full-length messenger RNA sequences greatly improve genome  
 annotation.";  
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY085804; AAM63020.1; -; mRNA.  
 DR Hypothetical protein.  
 SQ SEQUENCE 177 AA; 19819 MW; CC6A7841CBA8B38 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 177;  
 Best Local Similarity 25.0%; Pred. No. 2.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
 |  
 Db 87 CXXXXSSC 94

RESULT 13  
 Q9C9L7\_ARATH PRELIMINARY; PRT; 177 AA.  
 AC Q9C9L7;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein T6C23.4.  
 GN Name=T6C23.4;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,  
 RA Maiz R., Rensing C.M., Koo H., Fujii C.Y., Utterback T.R., C.M.,  
 RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Town C.D., Kaul S.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC013289; AAG52545.1; -; Genomic\_DNA.  
 KW PIR; F96719; F96719.  
 DR Hypothetical protein.  
 SQ SEQUENCE 177 AA; 19834 MW; 03D870608FEB8246 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 177;  
 Best Local Similarity 25.0%; Pred. No. 2.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
 |  
 Db 87 CXXXXSSC 94

RESULT 14  
 O60M47\_CABR PRELIMINARY; PRT; 199 AA.  
 ID O60M47;  
 AC O60M47;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein CBG3307.  
 GN Name=CBG3307;  
 OS Caenorhabditis briggsae.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderae; Caenorhabditis.  
 OC NCBI\_TaxID=6238;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA The C.briggsae Sequencing Consortium;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 2 BPTI/Kunitz inhibitor domains.  
 DR EMBL; CA001000141; CAB75329.1; -; Genomic\_DNA.  
 DR GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.  
 DR InterPro; IPR002223; Prot inh\_kunz-m.  
 DR InterPro; IPR006150; Worm repeat\_1.  
 DR Pfam; PF00014; Kunitz\_BPTI; 2.  
 DR PRINTS; PR00759; BASICPTASE.  
 DR ProDom; PD000222; Prot inh\_kunz-m; 2.  
 DR SMART; SM00131; KU; 2.  
 DR SMART; SM00289; WRI; 1.  
 DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
 DR PROSITE; PS0279; BPTI\_KUNITZ\_2; 2.

KW Hypothetical protein.  
 SQ SEQUENCE 199 AA; 21668 MW; 380A14BB90A152A3 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 199;  
 Best Local Similarity 25.0%; Pred. No. 3.1e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
 |  
 Db 112 CXXXXATTC 119

RESULT 15  
 O5VMP6\_ORISA PRELIMINARY; PRT; 200 AA.  
 ID O5VMP6;  
 AC O5VMP6;  
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Hypothetical protein OSUNBD008D07.38.  
 GN Name=OSUNBD008D07.38;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OC NCBI\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
 RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
 RA Hishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
 RA Ikano M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,  
 RA Kerasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
 RA Kuroki K., Maehara T., Mizuno H., Miyabayashi T., Mukai Y.,  
 RA Nagata H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshitara R., Yukawa K.,  
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.;  
 RT "The genome sequence and structure of rice chromosome 1.";  
 RL Nature 420:312-316(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP006237; BAB69279.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 200 AA; 23098 MW; 0E4249F62A3678AA CRC64;

Query Match 60.0%; Score 18; DB 2; Length 200;  
 Best Local Similarity 25.0%; Pred. No. 3.1e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 CXXXXXXC 11  
 |  
 Db 102 CXXXXATTC 109

RESULT 16  
 Q9EOL7\_MOUSE PRELIMINARY; PRT; 224 AA.  
 ID Q9EOL7;  
 AC Q9EOL7;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE TIS11D insertion variant (Fragment).  
 GN Name=Zfp3612; Synonym=Brf12;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

```

OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BLKS/J;
RA Cho K., Hobson K., Greenhalgh D.G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF205223; AAG45251.1; -; mRNA.
DR HSSP; P22893; 1M90.
DR SMR; Q9EOL7; 1-49.
DR MGI; MGI:107945; Zfp3612.
DR GO; GO:0003676; Functional acid binding; IEA.
DR InterPro; IPR000571; ZnF_CCH.
DR Pfam; PF00642; zf-CCH; 1.
DR SMART; SM00356; ZnF_CCH1; 1.
FT NON_TER 1
FT NON_TER 224
SQ SEQUENCE 224 AA; 23018 MW; 26E09C8465A5A61E CRC64;

Query Match
Best Local Similarity 25.0%; Score 18; DB 2; Length 224;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 129 CASSASSC 136

RESULT 17
Q6K2G5_ORYSA
ID Q6K2G5_ORYSA PRELIMINARY; PRT; 233 AA.
AC Q6K2G5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Glucanase-like.
GN Name=OSJNB0052M16.18;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 2, BAC
RT clone:OSJNB0052M16."
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005841; BAD23651.1; -; Genomic DNA.
DR Gramene; Q6K2G5; -.
SQ SEQUENCE 233 AA; 24907 MW; E4805DBE35FD0235 CRC64;

Query Match
Best Local Similarity 25.0%; Score 18; DB 2; Length 233;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 199 CAATRAAC 206

RESULT 18
OSJXD3_MAGR
ID OSJXD3_MAGR PRELIMINARY; PRT; 242 AA.
AC OSJXD3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Predicted protein.
GN ORFNames=MG07952.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.

```

```

OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Amburster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Bittsteyer B., Bloom T., Blye J., Boguslavsky L.,
RA Borowsky M., Boukhalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Chesnatsang Y., Citroen M.,
RA Collymore A., Condidine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degrey S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina F., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gairin G., Gierre S.,
RA Ghitre A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafer N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysseles M., Karlsson B.,
RA Kells C., Kieu A., Klesner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., MacDonald J., Maclean C., Major J.,
RA Manning J., Marabell R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrum J., Menues L.,
RA Mestrov J., Mhalale A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotocho B.,
RA O'Neill K., Osman S., Parker S., Perith D., Phunkhang P., Pigani B.,
RA Purcell S., Rachupka T., Ramasamy U., Raneau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schubach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sounguez C.,
RA Spencer B., Stalker J., Strange-thomson N., Scaavopoulos S.,
RA Stenson K., Stone S., Stone S., Stubbs M., Talamas J., Tchinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamta T., Tsomo N., Vallée D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Zander E.;
RT "The genome sequence of Magnaporthe grisea."
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01000977; EAA53675.1; -; Genomic DNA.
SQ SEQUENCE 242 AA; 25532 MW; 3339ABE816D10BA3 CRC64;

Query Match
Best Local Similarity 25.0%; Score 18; DB 2; Length 242;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 62 CASSSSC 69

RESULT 19
Q6BCP3_HPBVO
ID Q6BCP3_HPBVO PRELIMINARY; PRT; 245 AA.
AC Q6BCP3;

```

DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, last annotation update)  
 DE pre-S/S protein (Fragment).  
 OS Hepatitis B virus.  
 OC Virusae; Retro-transcribing viruses; Hepadnaviridae;  
 OC Orthohepadnavirus.  
 NC NCB1\_TaxID=10407;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Nguyen H.C., Dong S.H., Ho T.T.T., Ho H.T.D.;  
 RT "Sequence analysis of pre-S/S gene segment from some clinical  
 RT hepatitis B virus isolates.";  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY655569; AAT7983.1; -; Genomic\_DNA.  
 DR GO; GO:0003677; P.DNA binding; IEA.  
 DR GO; GO:0003887; P.DNA-directed DNA polymerase activity; IEA.  
 DR GO; GO:0006260; P.DNA replication; IEA.  
 DR GO; GO:0015032; P.viral life cycle; IEA.  
 DR InterPro; IPR000201; DNAPol\_viral\_N.  
 DR PANTHER; PTHR10832; Hepvir\_surfag; 1.  
 DR Pfam; PF00242; DNA\_pol\_viral\_N; 1.  
 DR Pfam; PF00695; vmsa; 1.  
 KM Antigen.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 245 AA; 26093 MW; C275D837A352298E CRC64;

Query Match 60.0%; Score 18; DB 2; Length 245;  
 Best Local Similarity 25.0%; Pred. No. 3.3e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 80 CASSTSSC 87

RESULT 20  
 Q9Y0E9 DROME PRELIMINARY; PRT; 258 AA.  
 ID Q9Y0E9  
 AC Q9Y0E9  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
 DE Salivary gland secretion protein (Fragment).  
 GN Name=Sgsl; Synonyms=Sgs-1; ORFNames=CG3047;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidae; Drosophilidae; Drosophila.  
 NC NCB1\_TaxID=7227;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Roth G.B., Matler S., Bornscheim H., Lehmann M., Korge G.;  
 RT "Structure and regulation of the salivary gland secretion protein gene  
 RT Sgs-1 of Drosophila melanogaster.";  
 RL Genetics 0:0-0(1999).  
 DR EMBL; AF156227; AAD43809.1; -; Genomic\_DNA.  
 DR FlyBase; FBgn0003372; CG3047.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 258 AA; 27832 MW; 0D9BF9218E6742EC CRC64;

Query Match 60.0%; Score 18; DB 2; Length 258;  
 Best Local Similarity 25.0%; Pred. No. 3.4e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 226 CTTTSSC 233

RESULT 21  
 Q7PR07 ANOGA PRELIMINARY; PRT; 269 AA.  
 ID Q7PR07  
 AC Q7PR07  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)  
 DE ENSANGP0000001657 (Fragment).  
 GN ORFNames=ENSANG0000001387;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;  
 OC Anophelinae; Anopheles.  
 NC NCB1\_TaxID=180454;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=PEST;  
 RG The Anopheles gambiae Sequence Committee;  
 RT "Anopheles gambiae re-annotation.";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=PEST;  
 RG The Anopheles gambiae Sequence Committee;  
 RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAB0100847; EAA06779.3; -; Genomic\_DNA.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 269 AA; 24228 MW; 25BBF71FD71F1F2 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 269;  
 Best Local Similarity 25.0%; Pred. No. 3.4e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 55 CSATSSC 62

RESULT 22  
 Q7Y198 ORYSA PRELIMINARY; PRT; 319 AA.  
 ID Q7Y198  
 AC Q7Y198  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, last annotation update)  
 DE Hypothetical protein OSUNBA0039018.17 (Hypothetical protein  
 DE OSUNB0036M02.4).  
 GN Name=OSUNBA0039018.17; Synonyms=OSUNB0036M02.4;  
 OS Oryza sativa (Japanese cultivar-group);  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoideae; Oryza.  
 NC NCB1\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,  
 RA Overton II L.L., Taitzin T., Kim M.M., Bera J.J., Jin S.S.,  
 RA Padrosch D.W., Tallon L.J., Koo H., Zigmund V., Heisao J., Blunt S.,  
 RA Vanaken S.S., Riedmuller S.B., Utecherback T.T., Feldlyum T.V.,  
 RA Yang Q.Q., Haas B.J., Sun B.B., Peterson J.J., Quackemush J.,  
 RA White O., Salzberg S.L., Fraser C.M.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.

RA Buell R.;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Jones K.M.,

RA Tallon L.J., Feldblyum T.V., Tsaltrin T., Bera J.J., Kim M.M., Jin S.,  
 RA Fadrosch D., Utung H., Overton II L.L., Reardon M., Weaver B.,  
 RA Johri S., Utebhag T.R., Pai G., Smith S., Mortman J., Haas B.J.,  
 RA Zhu W., Yang Q., Koo H., Zismann V., Hsiao J., Quackenbush J.,  
 RA White O., Salzberg S.L., Fraser C.M.,  
 RT "Oryza sativa chromosome 3 BAC OSJNB0036M02 genomic sequence.";  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Buell R.;  
 RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC133930; AAP4639.1; -; Genomic\_DNA.  
 DR EMBL: AC145388; AAV09142.1; -; Genomic\_DNA.  
 DR Gramene; Q7Y198; -;  
 DR GO: GO:0000151; C:ubiquitin ligase complex; IEA.  
 DR GO: GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO: GO:0008270; F:zinc ion binding; IEA.  
 DR GO: GO:0016567; F:protein ubiquitination; IEA.  
 DR InterPro: IPR001841; znf\_ring.  
 DR SMART: SM00184; RING; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 319 AA; 33316 MW; FD51B063E2B4AB4 CRC64;  
 Query Match 60.0%; Score 18; DB 2; Length 319;  
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 4 CXXXXXXC 11  
 DB 296 CAAAAAAC 303

RESULT 23  
 O61A5\_CAEER PRELIMINARY; PRT; 320 AA.  
 ID O61A5;  
 AC O61A5;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein CBG13750.  
 GN Name=CBG13750;  
 OS Caenorhabditis briggsae.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6238;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RG The C.briggsae Sequencing Consortium;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL: CAAC0100066; CAE68107.1; -; Genomic\_DNA.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0004930; F:G-protein coupled receptor activity; IEA.  
 DR InterPro: IPR000168; Nm7TM\_chemrecept.  
 KW Hypothetical protein.  
 SQ SEQUENCE 320 AA; 36037 MW; 5E3389F6C7CF3C7B CRC64;  
 Query Match 60.0%; Score 18; DB 2; Length 320;  
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 4 CXXXXXXC 11  
 DB 302 CASSTSTC 309

RESULT 24  
 O6ZNM3\_HUMAN PRELIMINARY; PRT; 328 AA.  
 ID O6ZNM3;  
 AC O6ZNM3;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein FLJ29006.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Stomach mucosa;  
 RA Nimmla K., Wagatsuna M., Kanda K., Kondo H., Yokoi T., Kodaira H.,  
 RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,  
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Suzuki Y., Hata H.,  
 RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,  
 RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,  
 RA Nagai K., Isogai T., Sugano S.,  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK131040; BAC85481.1; -; mRNA.  
 SQ SEQUENCE 328 AA; 35381 MW; 7D04FC08E3270F8D CRC64;  
 Query Match 60.0%; Score 18; DB 2; Length 328;  
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 4 CXXXXXXC 11  
 DB 4 CTSASSSC 11

RESULT 25  
 O6ZAB6\_ORYSA PRELIMINARY; PRT; 332 AA.  
 ID O6ZAB6\_ORYSA  
 AC O6ZAB6;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Limonene cyclase like protein.  
 GN Name=OSJNB0084107.20-2; Synonyms=OSJNB0002109.8-2;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSJNB0084107.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSJNB0002109.";  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP005179; BAC83914.1; -; Genomic\_DNA.  
 DR EMBL: AP005877; BAC31847.1; -; Genomic\_DNA.  
 DR Gramene; Q6Z486; -;  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR InterPro: IPR004263; Exostosin.  
 DR Pfam: PF03016; Exostosin; 1.  
 SQ SEQUENCE 332 AA; 34451 MW; 57CD6BBFC0CC5D6E CRC64;  
 Query Match 60.0%; Score 18; DB 2; Length 332;  
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 4 CXXXXXXC 11  
 DB 27 CAAAAAAC 34

```

RESULT 26
Q94589_9CILI PRELIMINARY; PRT; 350 AA.
ID Q94589_9CILI
AC Q94589
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Lemnabion bullinum.
OC Lemnabion bullinum.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
OC Lemnabion.
OC NCBI_TaxID=54108;
OX [1]
RP NUCLEOTIDE SEQUENCE.
RA Peters-Regehr T., Kusch J., Heckmann K.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y09220; CAA70420.1; -, Genomic_DNA.
DR HSSP; O16119; 1E2G.
DR InterPro; IPR006212; Furin_repeat.
DR SMART; SM00261; FU; 3.
SQ SEQUENCE 350 AA; 35159 MW; DBE0C67654B9D92E CRC64;

Query Match 60.0%; Score 18; DB 2; Length 350;
Beet Local Similarity 25.0%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 105 CSTSATTC 112

RESULT 27
RHOM DROME STANDARD; PRT; 355 AA.
ID P20350; Q9W0F2;
AC P20350; Q9W0F2;
DT 01-FEB-1991 (Rel. 17, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Rhomboid protein (EC 3.4.21.-) (Vainlet protein).
GN Name=rho; Synonyms=Ve; ORFNames=CG1004;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
OX [1]
RP NUCLEOTIDE SEQUENCE.
RA Bier E., Jan L.Y., Jan Y.N.;
RL MEDLINE=90249726; PubMed=2110920;
RT "Rhomboid, a gene required for dorsoventral axis establishment and
peripheral nervous system development in Drosophila melanogaster.";
RL Genes Dev. 4:190-203(1990).
[2]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=Berkely;
RC MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amarantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.C., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

```

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mikhina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy J., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacelb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskay R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach U.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[3]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Miura S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Betencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[4]
RP FUNCTION, AND MUTAGENESIS OF TRP-151; ARG-152; ASN-169; GLY-215;
RP SER-217 AND HIS-281.
RX MEDLINE=21526629; PubMed=11672525; DOI=10.1016/S0092-8674(01)00555-6;
RA Urban S., Lee J.R., Freeman M.;
RT "Drosophila Rhomboid-1 defines a family of putative intramembrane
serine proteases.";
RL Cell 107:173-182(2001).
CC -!- FUNCTION: Acts early in embryonic development to establish
position along the dorsoventral axis and then again later to
specify the fate of neuronal precursor cells. Involved in EGF
receptor signaling; cleaves Spitz to release the active growth
factor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Golgi.
CC -!- DEVELOPMENTAL STAGE: Early blastoderm stages and later during
neuron development.
CC -!- SIMILARITY: Belongs to the peptidase S54 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X52454; CAA36692.1; -, mRNA.
CC EMBL; AE003471; AAF47496.1; -, Genomic_DNA.
CC PIR; A34597; A34597.
CC MEROPS; S54.001;
CC Ensembl; CG1004; Drosophila melanogaster.
CC Flybase; FBgn0004635; rho.
CC GO; GO:0005794; C:Golgi apparatus; IDA.
CC GO; GO:0016021; C:integral to membrane; TAS.
CC GO; GO:0008236; P:serine-type peptidase activity; IDA.
CC GO; GO:0007420; P:brain development; IMP.
CC GO; GO:0046845; P:branched duct epithelial cell fate determin. .; TAS.
CC GO; GO:0001763; P:branching morphogenesis; NAS.
CC GO; GO:0007174; P:dermal morphogenesis; IGI.
CC GO; GO:0007479; P:leg disc proximal/distal pattern formation; TAS.
CC GO; GO:0007438; P:neocyte development; IGI.

```



DR GO; GO:0030707; P:ovarian follicle cell development (sensu In. . .; TAS.  
 DR GO; GO:0007422; P:peripheral nervous system development; TAS.  
 DR GO; GO:0045742; P:positive regulation of epidermal growth fac. . .; TAS.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IGI.  
 DR GO; GO:0007176; P:regulation of epidermal growth factor recep. . .; TAS.  
 DR GO; GO:0007432; P:salivary gland determination; NAS.  
 DR GO; GO:0035202; P:tracheal sac formation (sensu Insecta); TAS.  
 DR GO; GO:0007474; P:wing vein specification; NAS.  
 DR InterPro: IPR002610; Rhomboid\_Like.  
 DR Pfam: PF01694; Rhomboid; 1.  
 KW Developmental protein; Golgi stack; Hydrolyase; Protease;  
 KW Serine protease; Transmembrane.  
 FT TOPO\_DOM 1 98  
 FT TRANSMEM 99 119  
 FT TOPO\_DOM 120 162  
 FT TRANSMEM 163 183  
 FT TOPO\_DOM 184 188  
 FT TRANSMEM 189 209  
 FT TOPO\_DOM 210 210  
 FT TRANSMEM 211 231  
 FT TOPO\_DOM 232 244  
 FT TRANSMEM 245 265  
 FT TOPO\_DOM 266 275  
 FT TRANSMEM 276 296  
 FT TOPO\_DOM 297 308  
 FT TRANSMEM 309 329  
 FT TOPO\_DOM 330 355  
 FT ACT\_SITE 169 169  
 FT ACT\_SITE 217 217  
 FT ACT\_SITE 281 281  
 FT MUTAGEN 151 151  
 FT MUTAGEN 152 152  
 FT MUTAGEN 169 169  
 FT MUTAGEN 215 215  
 FT MUTAGEN 217 217  
 FT MUTAGEN 281 281  
 FT CONFLICT 4 4  
 FT CONFLICT 33 33  
 FT CONFLICT 46 46  
 FT CONFLICT 46 46  
 SQ SEQUENCE 355 AA; 39330 MW; 2E0572D08831C5D0 CRC64;  
 Query Match 60.0%; Score 18; DB 1; Length 355;  
 Best Local Similarity 25.0%; Pred. No. 3.8e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 4 CXXXXXXC 11  
 DB 57 CSTASSTC 64  
 RESULT 28  
 Q540V7 DROME  
 ID Q540V7 DROME PRELIMINARY; PRT; 355 AA.  
 AC Q540V7;  
 DT 13-SRP-2005 (TReMBLrel. 31, Created)  
 DT 13-SRP-2005 (TReMBLrel. 31, Last sequence update)  
 DT 13-SRP-2005 (TReMBLrel. 31, Last annotation update)  
 DE ID06131P.  
 GN Name=rho;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_Taxid=7227;  
 OK [1]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Berkeley.  
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nuno J., Paclel U., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celniker S.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY121641; AAM51968.1; -, mRNA.  
 SQ SEQUENCE 355 AA; 39330 MW; 2E0572D08831C5D0 CRC64;  
 Query Match 60.0%; Score 18; DB 2; Length 355;  
 Best Local Similarity 25.0%; Pred. No. 3.8e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 4 CXXXXXXC 11  
 DB 57 CSTASSTC 64  
 RESULT 29  
 T1SD MOUSE  
 ID T1SD MOUSE STANDARD; PRT; 367 AA.  
 AC P23949;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Butyrate response factor 2 (T1SID protein).  
 GN Name=zfp3612; Synonyms=Btf2, T1SID;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 OC NCBI\_Taxid=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=BALB/c;  
 RA MEDLINE=91141531; PubMed=1996120;  
 RA Varum B.C., Ma Q., Chi T., Fletcher B., Herachman H.R.;  
 RT "The T1sid primary response gene is a member of a gene family that  
 RT encodes proteins with a highly conserved sequence containing an  
 RT unusual Cys-His repeat.";  
 RT Mol. Cell. Biol. 11:1754-1758 (1991).  
 RL [2]  
 RP NUCLEOTIDE SEQUENCE OF 1-50.  
 RA Fletcher B.S.;  
 RL Thesis (1992); University of California Los Angeles, United States.  
 CC -1- FUNCTION: Probable regulatory protein involved in regulating the  
 CC response to growth factors.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -1- SIMILARITY: Contains 2 C3H1-type zinc fingers.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; M58564; AAA72946.1; -, mRNA.  
 DR EMBL; M97165; AAA97909.1; -, Genomic DNA.  
 DR PIR; C39590; C39590.  
 DR HSRP; P22893; 1M90.  
 DR SMR; P23949; 124-193.  
 DR Ensembl: ENSMUSG0000045817; Mus musculus.  
 DR MGI; MGI:107945; Zfp3612.  
 DR InterPro: IPR007635; T1SID\_N.  
 DR InterPro: IPR000571; ZnF\_CCH.  
 DR Pfam; PF04553; T1SID\_C; 1.  
 DR Pfam; PF00642; zf-CCH; 2.  
 DR SMART; SM00356; ZnF\_C3H1; 2.  
 KW DNA-binding; Metal-Binding; Nuclear protein; Repeat; Zinc;  
 KW Zinc-finger.  
 FT ZN\_FING 132 151  
 FT ZN\_FING 170 189  
 FT COMPBIAS 64 67  
 FT COMPBIAS 111 114  
 FT COMPBIAS 200 203  
 FT COMPBIAS 263 266  
 FT COMPBIAS 297 303  
 SQ SEQUENCE 367 AA; 37593 MW; 361244AF6244E46E CRC64;

Query Match 60.0%; Score 18; DB 1; Length 367;  
 Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 273 CSSSSSSC 280

## RESULT 30

ID Q86SB6\_DROVI PRELIMINARY; PRT; 404 AA.

AC Q86SB6; 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DE Rhoimoid.  
 GN Name=rho;  
 OS Drosophila virilis (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OK NCBI\_TaxID=7244;  
 RN [1]

RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22844358; PubMed=12963110; DOI=10.1016/S0925-4773(03)00164-3;  
 RA Nakamura Y., Matsuno K.;

RT "Species-specific activation of EGF receptor signaling underlies  
 evolutionary diversity in the dorsal appendage number of the genus  
 Drosophila eggshells.";  
 RL Mech. Dev. 120:897-907(2003).  
 DR EMBL; AB089248; BAC56701.1; -; mRNA.  
 DR MEROPS; S54.001; -;  
 DR FLYBASE; FBgn0062278; Dvir/rho.  
 DR InterPro; IPR002610; Rhomboid\_Like.  
 DR Pfam; PF01694; Rhomboid; 1.  
 SQ SEQUENCE 404 AA; 44035 MW; EF641632B02F9011 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 404;  
 Best Local Similarity 25.0%; Pred. No. 4e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 63 CSTRASSTC 70

## RESULT 31

ID Q6Z4B5\_ORYSA PRELIMINARY; PRT; 417 AA.

AC Q6Z4B5; 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Limonene cyclase like protein.  
 GN Name=OSUNB0084L07.20-3; Synonyms=OSUNB0002L09.8-3;  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Erihartoideae; Oryzaceae; Oryza.  
 OK NCBI\_TaxID=39947;

RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Katayose Y.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC  
 clone:OSUNB0084L07.";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Katayose Y.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC  
 clone:OSUNB0002L09.";

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP005179; BAC83915.1; -; Genomic\_DNA.  
 DR EMBL; AP005877; BAC31849.1; -; Genomic\_DNA.  
 DR Gramene; Q6Z4B5; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR004263; Exostein.  
 DR Pfam; PF03016; Exostein; 1.  
 SQ SEQUENCE 417 AA; 44610 MW; 67668E9A53B1DC6A CRC64;

Query Match 60.0%; Score 18; DB 2; Length 417;  
 Best Local Similarity 25.0%; Pred. No. 4.1e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 27 CAAAAAAC 34

## RESULT 32

ID Q8U0L4\_ORYLA PRELIMINARY; PRT; 445 AA.

AC Q8U0L4; 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DE TAPBP protein.  
 GN Name=TAPBP;  
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Perciformes; Atherinomorpha;  
 OC Belontiiformes; Adrianchthyidae; Oryziatidae; Oryzias.  
 OK NCBI\_TaxID=8090;  
 RN [1]

RP NUCLEOTIDE SEQUENCE.  
 RX STRAIN=Hd-IR;  
 RC MEDLINE=21850510; PubMed=11862394; DOI=10.1007/s00251-001-0427-3;  
 RA Matsuo M.Y., Asakawa S., Shimizu N., Kimura H., Nonaka M.;

RT "Nucleotide sequence of the MHC class I genomic region of a teleost,  
 the medaka (Oryzias latipes).";  
 RL Immunogenetics 53:930-940(2002).  
 DR EMBL; BA000027; BAB83851.1; -; Genomic\_DNA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0019885; P:antigen processing, endogenous antigen via . . .; IEA.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR008056; Tapasin.  
 DR Pfam; PF00047; Ig; 1.  
 DR PRINTS; PR01669; TAPASIN.  
 DR SMART; SM00409; Ig; 2.  
 DR PROSITE; PS50835; IG\_LIKE; 2.  
 DR KEGG; K01101; Immunoglobulin domain; Repeat.  
 SQ SEQUENCE 445 AA; 48260 MW; FDF7C5FBE74370BB CRC64;

Query Match 60.0%; Score 18; DB 2; Length 445;  
 Best Local Similarity 25.0%; Pred. No. 4.2e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11  
 DB 22 CSSSSSSC 29

## RESULT 33

ID Q7SHC6\_NEUCR PRELIMINARY; PRT; 480 AA.

AC Q7SHC6; 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Predicted protein (Hypothetical protein B13N4.210).  
 GN Name=NCU01880.1; Synonyms=B13N4.210;  
 OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 NCBI\_TaxID=5141;  
 [1]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=0874A;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rahman B.,  
 RA Blinks T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Iankilev P., Pedersen D., Nelson M., Washburne M.,  
 RA Selicremikoff C.F., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
 RA Kamal M., Kamysseilis M., Mauceli E., Bielke C., Rudd S., Frishman D.,  
 RA Kryzofowa S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,  
 RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,  
 RA Desouza C.C., Glas L., Orbach M.J., Berglund J., Voelker R.,  
 RA Yarden O., Plamann M., Siller S., Dunlap J., Radford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbold D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;  
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";  
 RL Nature 0:0-0(2003).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Schulte U., Aign V., Hobeisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RA German Neurospora genome project;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AABX01000004; EAA36275.1; -, Genomic DNA.  
 DR EMBL; BX842681; CAE81980.1; -, Genomic DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 480 AA; 50449 MW; E073DB8EC33C8A2B CRC64;  
 QY 4 CXXXXXXC 11  
 Db 127 CSSSSSSC 134  
 Query Match 60.0%; Score 18; DB 2; Length 480;  
 Best Local Similarity 25.0%; Pred. No. 4.3e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

DR GO; GO:0004114; F.3',5'-cyclic-nucleotide phosphodiesterase a. .; IEA.  
 DR GO; GO:0003824; Fcatalytic activity; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR003607; Met\_phos\_hydro.  
 DR InterPro; IPR002073; PDBase.  
 DR Pfam; PF00233; PDBase\_1; 1.  
 DR PRINTS; PR00387; PDIESTERASB1.  
 DR SMART; SM00471; Hdc; 1.  
 DR PROSITE; PS00126; PDBASE\_1; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 480 AA; 53835 MW; 75AB3894A55D01CB CRC64;  
 QY 4 CXXXXXXC 11  
 Db 205 CTAAATAC 212  
 Query Match 60.0%; Score 18; DB 2; Length 480;  
 Best Local Similarity 25.0%; Pred. No. 4.3e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

RESULT 35  
 T1SD\_HUMAN  
 ID T1SD\_HUMAN STANDARD; PRT; 492 AA.  
 AC P47974; Q9BSJ3;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Butyrate response factor 2 (T1SD1D protein) (BGF-response factor 2)  
 DE (BRF-2).  
 GN Name=ZFP3612; Synonyms=BRF2, BRF2, T1SD1D;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RX MEDLINE=96132724; PubMed=8545129;  
 RP NUCLEOTIDE SEQUENCE.  
 RA Ino T., Yasui H., Hirano M., Kurosawa Y.;  
 RT "Identification of a member of the T1SD1 early response gene family at  
 RT the insertion point of a DNA fragment containing a gene for the T-cell  
 RT receptor beta chain in an acute T-cell leukemia.";  
 RL Oncogene 11:2705-2710(1995).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=95137407; PubMed=7835719; DOI=10.1016/0378-1119(94)00696-P;  
 RA Nle X.F., Maclean K.N., Kumar V., McKay I.A., Bustin S.A.;  
 RT "ERF-2, the human homologue of the murine T1sd1 early response  
 RT gene.";  
 RL Gene 152:285-286(1995).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Haile F.,  
 RA Stachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Umed T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Lotteliano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Botterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.B.,  
 RA Scherch A., Schein J.E., Jones S.J.W., Marra W.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human

```

RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Probable regulatory protein involved in regulating the
CC response to growth factors.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Contains 2 C3H1-type zinc fingers.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
CC EMBL; U07802; AAA1778.1; -; Genomic_DNA.
DR EMBL; X78992; CAAS5592.1; -; mRNA.
DR EMBL; BC005010; AAH05010.1; -; mRNA.
DR PIR; S49147; S49147.
DR PDB; 1RGO; NMR; A=151-220.
DR Ensemble; ENSG00000152518; Homo sapiens.
DR HGNC; HGNC:1108; ZFP36L2.
DR H-InvDB; HIX002007; -.
DR GO; GO:0003700; F:transcription factor activity; TAS.
DR GO; GO:0008283; P:cell proliferation; TAS.
DR InterPro; IPR007635; T1a1b.N.
DR InterPro; IPR000571; ZnF_CCH.
DR Pfam; PF04553; T1a1b_C; 1.
DR Pfam; PF00642; zf-CCH1; 2.
DR SMART; SM00356; ZnF_C3H1; 2.
DR 3D-structure; DNA-binding; Metal-binding; Nuclear protein; Repeat;
KW Zinc; zinc-finger.
FT ZN_FING 159 178 C3H1-type 1.
FT ZN_FING 197 216 C3H1-type 2.
FT COMPBIAS 106 109 Poly-Gly.
FT COMPBIAS 138 141 Poly-Gln.
FT COMPBIAS 143 146 Poly-Gly.
FT COMPBIAS 288 291 Poly-Pro.
FT COMPBIAS 323 330 Poly-Ala.
FT COMPBIAS 382 388 Poly-Ala.
FT COMPBIAS 393 399 Poly-Gln.
FT COMPBIAS 96 97 TS -> DL (in Ref. 1).
FT COMPBIAS 318 318 A -> T (in Ref. 2).
FT COMPBIAS 329 330 AA -> LR (in Ref. 2).
FT COMPBIAS 330 330 A -> AAA (in Ref. 3).
FT COMPBIAS 399 399 Q -> QQQQ (in Ref. 3).
FT COMPBIAS 451 460 Missing (in Ref. 1).
FT COMPBIAS 492 AA; 50921 MW; E4E2EE26791CCP7 CRC64;
SQ SEQUENCE

Query Match 60.0%; Score 18; DB 1; Length 492;
Best Local Similarity 25.0%; Pred. No. 4,4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

RT "The sequence of Homo sapiens BAC clone RP11-339H12."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RL Waterston R.H.;
RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC010883; AY14992.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 494 AA; 51062 MW; 10E23FA2BDABD4 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 494;
Best Local Similarity 25.0%; Pred. No. 4,4e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```
DR GO; GO:0005070; F:SH3/SH2 adaptor activity; TAS.
DR GO; GO:0042100; P:B-cell proliferation; IDA.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
FT NON TER
SQ SEQUENCE 526 AA; 57393 MW; 6F358E931E2C63E0 CRC64;

Query Match
Best Local Similarity 25.0%; Score 18; DB 2; Length 526;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 71 CSASASAC 78

RESULT 38
Q5B195 DROME PRELIMINARY; PRT; 544 AA.
ID Q5B195 DROME PRELIMINARY; PRT; 544 AA.
AC Q5B195;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE RE04051P.
GN Name=dpa;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkely;
RA Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA Park S., Wan K., Yu C., Rubin G.M., Celinker S.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT021329; AAX347.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008094; F:DNA replication initiation; IEA.
DR GO; GO:0006270; P:DNA replication initiation; IEA.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR001208; MCM.
DR InterPro; IPR008047; MCM 4.
DR PRINTS; PR01660; MCMPROTEIN4.
DR SMART; SM00350; MCM; 1.
DR PROSITE; PS00615; C:TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE; PS50051; MCM 2; 1.
SQ SEQUENCE 544 AA; 60347 MW; 8D221F998960D707 CRC64;

Query Match
Best Local Similarity 25.0%; Score 18; DB 2; Length 544;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 522 CSTRSTSC 529

RESULT 39
Q91TR2 TUVH1 PRELIMINARY; PRT; 545 AA.
ID Q91TR2 TUVH1 PRELIMINARY; PRT; 545 AA.
AC Q91TR2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T31.
Tupaid herpesvirus 1 (strain 1) (TUVH-1) (Herpesvirus tupai (strain
```

```
OS 1)).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10397;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2;
RX MEDLINE=21211637; PubMed=11312357;
RX DOI=10.1126/JVI.75.10.4854-4870.2001;
RA Bahr U., Darai G.;
RT "Analysis and characterization of the complete genome of tupai (tree
shrew) herpesvirus.";
RL J. Virol. 75:4854-4870(2001).
DR EMBL; AF281817; AAK57075.1; -; Genomic_DNA.
DR InterPro; IPR007578; DUF570.
DR Pfam; PF04489; DUF570; 1.
SQ SEQUENCE 545 AA; 59508 MW; 98B5ECDEFF7A806 CRC64;

Query Match
Best Local Similarity 25.0%; Score 18; DB 2; Length 545;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 535 CAASATC 542

RESULT 40
Q528N9 ORYSA PRELIMINARY; PRT; 551 AA.
ID Q528N9 ORYSA PRELIMINARY; PRT; 551 AA.
AC Q528N9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Aspartic proteinase nepenthesin II-like.
GN Name=P0541C02.19-2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
clone:P0541C02.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003769; BAD61723.1; -; Genomic DNA.
DR GO; GO:0004194; F:pepsin A activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001461; Peptidase_A1.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00026; Asp; 1.
DR PRINTS; PR00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
SQ SEQUENCE 551 AA; 57586 MW; BE78840EB14824DF CRC64;

Query Match
Best Local Similarity 25.0%; Score 18; DB 2; Length 551;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 182 CATATSSC 189

RESULT 41
Q8SB30 ORYSA PRELIMINARY; PRT; 551 AA.
ID Q8SB30 ORYSA PRELIMINARY; PRT; 551 AA.
AC Q8SB30;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
```

```

DE Putative chloroplast nucleoid DNA-binding protein.
GN Name=OJ1540_H01.13;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=4530;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan O., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Ganabinger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Teitelin T., Riggs F., Hejao J., Zisman V., Blunt S., Pai G.,
RA Vanden S.B., Utterback T.R., Feldlyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC091774; AAL79734.1; -; Genomic_DNA.
DR HSSP; P00797; 2REN.
DR Gramene; O8SB30; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004194; F:pepsin A activity; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR01461; Peptidase_A1.
DR InterPro; IPR01969; Pept_Asp_AS.
DR Pfam; PF00026; Asp_1.
DR PRINTS; PR00797; PEP5IN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
DR DNA-binding.
SQ SEQUENCE 551 AA; 57586 MW; BE78840EB14824DF CRC64;

Query Match 60.0%; Score 18; DB 2; Length 551;
Best Local Similarity 25.0%; Pred. No. 4.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 182 CATATSSC 189

RESULT 42
Q5AP57 CANAL PRELIMINARY; PRT; 552 AA.
AC Q5AP57;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical YFW family protein 5.
GN ORFNames=C6019.4881;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OC NCBI_TaxID=337561;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=SC5314;
RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RT "The diploid genome sequence of Candida albicans."
RT Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=SC5314;
RA Dungan J., Kuo A., Newport G., Ian C.-Y., Iijima C., Adegoke O.,
RA Roberts J., Persson K., Donnelly S., Favoretto S., Tsung K.-W.,
RA Jones T., Scherer S., Agabian N.;
RT "Annotation of the Genome of Candida albicans."
RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an

```

```

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAC001000001; EAL04869.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 552 AA; 62256 MW; 5CA72B0ED5B87B3F CRC64;

Query Match 60.0%; Score 18; DB 2; Length 552;
Best Local Similarity 25.0%; Pred. No. 4.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 466 CTSATTSC 473

RESULT 43
O80J53_HPBV0 PRELIMINARY; PRT; 560 AA.
AC O80J53;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Truncated polymerase.
OC Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lin X., Xu X., Zheng D.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY238972; ANO64453.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003644; F:RNA-directed DNA polymerase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:006278; F:RNA-dependent DNA replication; IEA.
DR InterPro; IPR001462; DNAPol_viral_C.
DR InterPro; IPR000201; DNAPol_viral_N.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00336; DNA_pol_viral_C_1.
DR Pfam; PF00242; DNA_pol_viral_N_1.
DR Pfam; PF00078; RVT_1_1.
SQ SEQUENCE 560 AA; 62448 MW; 7EC0D0F6F991CE4B CRC64;

Query Match 60.0%; Score 18; DB 2; Length 560;
Best Local Similarity 25.0%; Pred. No. 4.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CXXXXXXC 11
DB 268 CASSSSSC 275

RESULT 44
Q6N1S1_CORDI PRELIMINARY; PRT; 581 AA.
AC Q6N1S1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative secreted protein.
GN Ordered locus Names=DIP0696;
OC Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=Blotcyne graves / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gk9874;
RA Cerdano-Tarraga A.-M., Estratton A., Dover L.G., Holden M.T.G.,

```



```

RESULT 47
O4RVN9_TETNG PRELIMINARY; PRT; 691 AA.
ID O4RVN9_TETNG PRELIMINARY; PRT; 691 AA.
AC O4RVN9;
13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 10 SCAP15009, whole genome shotgun sequence.
GN ORFname=GSTENG00031390001;
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_TaxId=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Coataz C., Bernot A.,
RA Nicard S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Sainoubat M., Levy M., Boudet N., Castellano S.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Crnaud C., Duprat S., Broctier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lander V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.,
RA "Genome duplication in the teleost fish Tetradodon nigroviridis reveals
RT the early vertebrate proco-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAE01015009; CAG09923.1; -; Genomic DNA.
SQ SEQUENCE 691 AA; 78163 MW; 440CAE5E405255CD CRC64;

Query Match 60.0%; Score 18; DB 2; Length 691;
Best Local Similarity 25.0%; Pred. No. 5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 CXXXXXXC 11
DB 288 CSSSSSSC 295

RESULT 48
O7S136_NEUCR PRELIMINARY; PRT; 705 AA.
ID O7S136_NEUCR PRELIMINARY; PRT; 705 AA.
AC O7S136;
01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU09990.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
NCBI_TaxId=5141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=OR74A;
RC STRAIN=OR74A;
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Federsen D., Nelson M., Washburne M.,

```

```

RA Selltremlkoff C.P., Kinsey J.A., Braut E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamyssellis M., Mauceli E., Bieleke C., Rudd S., Frishman D.,
RA Kryetofova S., Raamussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,
RA Desouza C.C., Glas J., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plmann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannheim G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000465; EAA29063.1; -; Genomic DNA.
SQ SEQUENCE 705 AA; 75444 MW; D7AB40BF0243E134 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 705;
Best Local Similarity 25.0%; Pred. No. 5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 CXXXXXXC 11
DB 367 CTSSTATAC 374

RESULT 49
O5RFW2_BRARE PRELIMINARY; PRT; 711 AA.
ID O5RFW2_BRARE PRELIMINARY; PRT; 711 AA.
AC O5RFW2;
01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Novel protein (Zgc:77446).
GN Name=CH211-200P13.3; ORFNames=CH211-200P13.3-001, zgc:77446;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
NCBI_TaxId=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Donaldson S.,
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CHS36610; CA11747.1; -; Genomic DNA.
DR ZFIN; ZDB-GENE-040426-933; zgc:77446.
DR InterPro; IPR003892; CUE.
DR InterPro; IPR001876; Znf_RangDP.
DR Pfam; PF02845; CUE; 1.
DR SMART; SM00546; CUE; 1.
DR SMART; SM00547; Znf_RBZ; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS0199; ZF_RANBP2_2; 1.
SQ SEQUENCE 711 AA; 76880 MW; E26723DFF30859356 CRC64;

Query Match 60.0%; Score 18; DB 2; Length 711;
Best Local Similarity 25.0%; Pred. No. 5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 4 CXXXXXXC 11
DB 330 C SSTSSC 337

RESULT 50
O5AP84_CANAL PRELIMINARY; PRT; 722 AA.
ID O5AP84_CANAL PRELIMINARY; PRT; 722 AA.
AC O5AP84;
10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical YFM family protein 5.

```



GN Name=YFW5; ORFNames=CaO19.12344;  
OS Candida albicans SC5314.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
OX NCBI\_TaxID=237561;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=SC5314;  
RX PubMed=15123810; DOI=10.1073/pnas.0401648101;  
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,  
RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,  
RA Davis R.W., Scherer S.;  
RT "The diploid genome sequence of Candida albicans.";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=SC5314;  
RA Dungan J., Kuo A., Newport G., Ian C.-Y., Iijima C., Adegbola O.,  
RA Roberts J., Persson K., Donnelly S., Favoreto S., Tung K.-W.,  
RA Jones T., Scherer S., Agabian N.;  
RT "Annotation of the Genome of Candida albicans.";  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AAC001000002; EAL04673.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 722 AA; 81770 MW; 16E7C88A35A416AE CRC64;  
  
Query Match 60.0%; Score 18; DB 2; Length 722;  
Best local Similarity 25.0%; Pred. No. 5.1e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 4 CXXXXXXC 11  
DB 636 CTSATTSC 643

Search completed: January 4, 2006, 16:08:27  
Job time : 124.426 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 : Search time 102.652 Seconds  
(without alignments)  
64.204 Million cell updates/sec

Title: US-09-932-322-3  
Perfect score: 31  
Sequence: 1 XXXXXXXXXXXXX 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 1000 summaries

Database : A\_Geneseq\_21.\*  
1: geneseq1980s:\*  
2: geneseq1990s:\*  
3: geneseq2000s:\*  
4: geneseq2001s:\*  
5: geneseq2002s:\*  
6: geneseq2003as:\*  
7: geneseq2003bs:\*  
8: geneseq2004s:\*  
9: geneseq2005s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysts of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	58.1	59	4	ABBA1147 Peptide #
2	18	58.1	59	4	AAAG4923 Peptide #
3	18	58.1	59	4	ABB25185 Protein #
4	18	58.1	59	4	AAW74807 Human bon
5	18	58.1	59	4	AAW62003 Human bra
6	18	58.1	59	4	ABG56589 Human liv
7	18	58.1	59	5	ABG44598 Human pep
8	18	58.1	108	3	AAAB2190 Exo14 par
9	18	58.1	134	7	ABO75321 Pseudom
10	18	58.1	139	7	ABO75320 Pseudom
11	18	58.1	139	7	ABO75320 Pseudom
12	18	58.1	140	7	AAW93094 Human end
13	18	58.1	140	4	AAU20102 Human dig
14	18	58.1	140	5	ABP40963 Human liv
15	18	58.1	140	7	ADJ15081 Human liv
16	18	58.1	144	8	ADP31453 Human sec
17	18	58.1	145	7	ABO83270 Pseudom
18	18	58.1	147	7	ABO79762 Pseudom
19	18	58.1	153	8	ADP31511 Human sec
20	18	58.1	155	8	ADN23733 Bacterial
21	18	58.1	156	7	ABO83103 Pseudom
22	18	58.1	156	8	ADP30759 Human sec
23	18	58.1	165	8	ADP31286 Human sec
24	18	58.1	165	8	ADP31174 Human sec

25	18	58.1	169	7	ABO83337 Pseudom
26	18	58.1	174	8	ADP30801 Human sec
27	18	58.1	193	7	ABO73977 Pseudom
28	18	58.1	198	8	ADP30493 Human sec
29	18	58.1	198	8	ADP30492 Human sec
30	18	58.1	198	8	ADP30477 Human sec
31	18	58.1	198	8	ADP30481 Human sec
32	18	58.1	198	8	ADP30491 Human sec
33	18	58.1	228	8	ADP31281 Human sec
34	18	58.1	229	7	ABO80501 Pseudom
35	18	58.1	233	7	ABO77555 Pseudom
36	18	58.1	234	8	ADP31468 Human sec
37	18	58.1	242	8	ADP31616 Human sec
38	18	58.1	246	5	AAU75930 Androgen
39	18	58.1	246	7	ABO74848 Pseudom
40	18	58.1	249	8	ADP30754 Human sec
41	18	58.1	254	8	ADP31336 Human sec
42	18	58.1	279	8	ADP30846 Human sec
43	18	58.1	279	8	ADP31523 Human sec
44	18	58.1	279	8	ADP30844 Human sec
45	18	58.1	281	4	AAW48050 Signal tr
46	18	58.1	281	8	ADW43729 Corn DREB
47	18	58.1	286	7	ABO79000 Pseudom
48	18	58.1	297	8	ADP31192 Human sec
49	18	58.1	312	8	ADP31454 Human sec
50	18	58.1	317	8	ADP31615 Human sec
51	18	58.1	339	8	ADP30519 Human sec
52	18	58.1	346	8	ADP31355 Human sec
53	18	58.1	348	8	ADP31441 Human sec
54	18	58.1	348	8	ADP30985 Human sec
55	18	58.1	348	8	ADP31676 Human sec
56	18	58.1	354	8	ADP30783 Human sec
57	18	58.1	354	8	ADP30779 Human sec
58	18	58.1	354	8	ADP30784 Human sec
59	18	58.1	354	8	ADP30778 Human sec
60	18	58.1	360	8	ADP31209 Human sec
61	18	58.1	363	8	ADP30659 Human sec
62	18	58.1	369	8	ADP30582 Human sec
63	18	58.1	372	8	ADP31492 Human sec
64	18	58.1	372	8	ADP31610 Human sec
65	18	58.1	382	8	ADP31221 Human sec
66	18	58.1	387	8	ADP31368 Human sec
67	18	58.1	396	8	ADP31378 Human sec
68	18	58.1	398	7	ABO71161 Pseudom
69	18	58.1	405	8	ADP31534 Human sec
70	18	58.1	411	8	ADP30729 Human sec
71	18	58.1	426	8	ADP31495 Human sec
72	18	58.1	442	8	ADP31134 Human sec
73	18	58.1	444	8	ADP31640 Human sec
74	18	58.1	471	8	ADP30854 Human sec
75	18	58.1	474	8	ADP31083 Human sec
76	18	58.1	483	8	ADP30852 Human sec
77	18	58.1	483	8	ADP30853 Human sec
78	18	58.1	484	8	ADP31628 Human sec
79	18	58.1	489	8	ADP31380 Human sec
80	18	58.1	519	8	ADP31190 Human sec
81	18	58.1	525	8	ADP31227 Human sec
82	18	58.1	528	8	ADP30529 Human sec
83	18	58.1	528	8	ADP31525 Human sec
84	18	58.1	533	8	ADP31279 Human sec
85	18	58.1	543	8	ADP30872 Human sec
86	18	58.1	549	8	ADP30855 Human sec
87	18	58.1	552	8	ADP31524 Human sec
88	18	58.1	558	8	ADP31254 Human sec
89	18	58.1	558	8	ADP31255 Human sec
90	18	58.1	592	8	ADP30917 Human sec
91	18	58.1	603	8	ADP31645 Human sec
92	18	58.1	610	8	ADP31264 Human sec
93	18	58.1	615	8	ADP31169 Human sec
94	18	58.1	615	8	ADP31132 Human sec
95	18	58.1	615	8	ADP31361 Human sec
96	18	58.1	615	8	ADP30803 Human sec
97	18	58.1	615	8	ADP31360 Human sec

98	18	58.1	637	8	ADp31397	Adp31397	Human	sec	171	18	58.1	1092	8	ADP31358	Adp31358	Human	sec
99	18	58.1	639	8	ADP30861	Adp30861	Human	sec	172	18	58.1	1107	8	ADP30504	Adp30504	Human	sec
100	18	58.1	642	9	ADz10464	Adz10464	P. gingiv		173	18	1110	8	ADP31430	Adp31430	Human	sec	
101	18	58.1	642	8	ADp30748	Adp30748	Human	sec	174	18	58.1	1116	8	ADP31692	Adp31692	Human	sec
102	18	58.1	660	8	ADP31606	Adp31606	Human	sec	175	18	58.1	1128	8	ADP30483	Adp30483	Human	sec
103	18	58.1	665	8	ADP30571	Adp30571	Human	sec	176	18	58.1	1133	8	ADP30931	Adp30931	Human	sec
104	18	58.1	681	8	ADP30868	Adp30868	Human	sec	177	18	58.1	1133	8	ADP30928	Adp30928	Human	sec
105	18	58.1	681	8	ADP31582	Adp31582	Human	sec	178	18	58.1	1133	8	ADP30933	Adp30933	Human	sec
106	18	58.1	687	8	ADP31364	Adp31364	Human	sec	179	18	58.1	1133	8	ADP30939	Adp30939	Human	sec
107	18	58.1	699	8	ADP30747	Adp30747	Human	sec	180	18	58.1	1165	8	ADP30879	Adp30879	Human	sec
108	18	58.1	705	8	ADP31623	Adp31623	Human	sec	181	18	58.1	1167	8	ADP30502	Adp30502	Human	sec
109	18	58.1	708	8	ADP31191	Adp31191	Human	sec	182	18	58.1	1168	8	ADP31046	Adp31046	Human	sec
110	18	58.1	711	8	ADP31440	Adp31440	Human	sec	183	18	58.1	1170	8	ADP30919	Adp30919	Human	sec
111	18	58.1	729	8	ADP30541	Adp30541	Human	sec	184	18	58.1	1170	8	ADP30922	Adp30922	Human	sec
112	18	58.1	732	8	ADP30650	Adp30650	Human	sec	185	18	58.1	1184	7	ADDP6553	Addp6553	Human	Ckb
113	18	58.1	735	8	ADP31010	Adp31010	Human	sec	186	18	58.1	1189	8	ADP31043	Adp31043	Human	sec
114	18	58.1	738	8	ADP31288	Adp31288	Human	sec	187	18	58.1	1191	8	ADP30993	Adp30993	Human	sec
115	18	58.1	750	8	ADP30857	Adp30857	Human	sec	188	18	58.1	1192	8	ADP30964	Adp30964	Human	sec
116	18	58.1	750	8	ADP31131	Adp31131	Human	sec	189	18	58.1	1199	8	ADP31044	Adp31044	Human	sec
117	18	58.1	756	8	ADP30856	Adp30856	Human	sec	190	18	58.1	1200	8	ADP30935	Adp30935	Human	sec
118	18	58.1	756	8	ADP31039	Adp31039	Human	sec	191	18	58.1	1205	8	ADP30898	Adp30898	Human	sec
119	18	58.1	762	8	ADP31012	Adp31012	Human	sec	192	18	58.1	1215	8	ADP30900	Adp30900	Human	sec
120	18	58.1	768	8	ADP30912	Adp30912	Human	sec	193	18	58.1	1215	8	ADP31293	Adp31293	Human	sec
121	18	58.1	768	8	ADP31362	Adp31362	Human	sec	194	18	58.1	1221	8	ADP31426	Adp31426	Human	sec
122	18	58.1	768	8	ADP31363	Adp31363	Human	sec	195	18	58.1	1224	8	ADP31426	Adp31426	Human	sec
123	18	58.1	771	8	ADP31469	Adp31469	Human	sec	196	18	58.1	1227	8	ADP31210	Adp31210	Human	sec
124	18	58.1	774	8	ADP30885	Adp30885	Human	sec	197	18	58.1	1233	8	ADP30523	Adp30523	Human	sec
125	18	58.1	780	8	ADP31077	Adp31077	Human	sec	198	18	58.1	1237	2	AAr313791	AAr313791	E75A	prot
126	18	58.1	795	8	ADP31332	Adp31332	Human	sec	199	18	58.1	1237	6	AAE30117	AAe30117	Fruit	fly
127	18	58.1	816	8	ADP31245	Adp31245	Human	sec	200	18	58.1	1260	8	ADP31533	Adp31533	Human	sec
128	18	58.1	816	8	ADP31581	Adp31581	Human	sec	201	18	58.1	1284	4	ABB71869	ABb71869	Drosophi	
129	18	58.1	828	8	ADP30878	Adp30878	Human	sec	202	18	58.1	1284	5	ABB08376	ABb08376	D. melano	
130	18	58.1	834	8	ADP30738	Adp30738	Human	sec	203	18	58.1	1302	8	ADP31695	Adp31695	Human	sec
131	18	58.1	849	8	ADP31182	Adp31182	Human	sec	204	18	58.1	1314	8	ADP31197	Adp31197	Human	sec
132	18	58.1	849	8	ADP31184	Adp31184	Human	sec	205	18	58.1	1339	8	ADP30732	Adp30732	Human	sec
133	18	58.1	849	8	ADP31020	Adp31020	Human	sec	206	18	58.1	1341	8	ADP30888	Adp30888	Human	sec
134	18	58.1	861	8	ADP31021	Adp31021	Human	sec	207	18	58.1	1344	8	ADP31211	Adp31211	Human	sec
135	18	58.1	865	8	ADP30566	Adp30566	Human	sec	208	18	58.1	1350	8	ADP31138	Adp31138	Human	sec
136	18	58.1	885	8	ADP31011	Adp31011	Human	sec	209	18	58.1	1358	8	ADP30995	Adp30995	Human	sec
137	18	58.1	885	8	ADP31127	Adp31127	Human	sec	210	18	58.1	1362	8	ADP31181	Adp31181	Human	sec
138	18	58.1	888	8	ADP30971	Adp30971	Human	sec	211	18	58.1	1365	8	ADP31035	Adp31035	Human	sec
139	18	58.1	888	8	ADP31648	Adp31648	Human	sec	212	18	58.1	1371	8	ADP31646	Adp31646	Human	sec
140	18	58.1	890	8	ADP31059	Adp31059	Human	sec	213	18	58.1	1371	8	ADP30876	Adp30876	Human	sec
141	18	58.1	892	8	ADP31578	Adp31578	Human	sec	214	18	58.1	1380	8	ADP31566	Adp31566	Human	sec
142	18	58.1	906	9	AEBS3801	AEb53801	Drosophi		215	18	58.1	1387	8	ADP30946	Adp30946	Human	sec
143	18	58.1	921	8	ADP31522	Adp31522	Human	sec	216	18	58.1	1417	8	ADP31160	Adp31160	Human	sec
144	18	58.1	930	8	ADP31444	Adp31444	Human	sec	217	18	58.1	1431	8	ADP31609	Adp31609	Human	sec
145	18	58.1	936	8	ADP31597	Adp31597	Human	sec	218	18	58.1	1435	6	ABU09754	ABu09754	Novel	hum
146	18	58.1	939	8	ADP30726	Adp30726	Human	sec	219	18	58.1	1435	6	ABU90136	ABu90136	Novel	hum
147	18	58.1	945	8	ADP31238	Adp31238	Human	sec	220	18	58.1	1435	6	ABU96438	ABu96438	Novel	hum
148	18	58.1	945	8	ADP31237	Adp31237	Human	sec	221	18	58.1	1435	6	ABU99047	ABu99047	Novel	hum
149	18	58.1	948	8	ADP30586	Adp30586	Human	sec	222	18	58.1	1435	6	ABU98262	ABu98262	Novel	hum
150	18	58.1	951	8	ADP30943	Adp30943	Human	sec	223	18	58.1	1435	6	ABU91968	ABu91968	Novel	hum
151	18	58.1	960	8	ADP31470	Adp31471	Human	sec	224	18	58.1	1435	6	ABU85272	ABu85272	Novel	hum
152	18	58.1	960	8	ADP31470	Adp31470	Human	sec	225	18	58.1	1435	6	ABO00411	ABo00411	Novel	hum
153	18	58.1	966	8	ADP30745	Adp30745	Human	sec	226	18	58.1	1435	6	ABU88962	ABu88962	Novel	hum
154	18	58.1	990	8	ADP31553	Adp31553	Human	sec	227	18	58.1	1435	6	ABO06458	ABo06458	Novel	hum
155	18	58.1	1002	8	ADP30866	Adp30866	Human	sec	228	18	58.1	1435	6	ABU95518	ABu95518	Novel	hum
156	18	58.1	1023	8	ADP30994	Adp30994	Human	sec	229	18	58.1	1435	6	ABU95208	ABu95208	Novel	hum
157	18	58.1	1032	8	ADP30874	Adp30874	Human	sec	230	18	58.1	1435	6	ABU90756	ABu90756	Novel	hum
158	18	58.1	1032	8	ADP30863	Adp30863	Human	sec	231	18	58.1	1435	6	ABU93918	ABu93918	Novel	hum
159	18	58.1	1033	8	ADP30984	Adp30984	Human	sec	232	18	58.1	1435	6	ABU86192	ABu86192	Novel	hum
160	18	58.1	1035	8	ADP31552	Adp31552	Human	sec	233	18	58.1	1435	6	ABU82047	ABu82047	Novel	hum
161	18	58.1	1044	8	ADP31517	Adp31517	Human	sec	234	18	58.1	1435	6	ABU07908	ABu07908	Novel	hum
162	18	58.1	1056	8	ADP31592	Adp31592	Human	sec	235	18	58.1	1435	6	ABU94228	ABu94228	Novel	hum
163	18	58.1	1057	8	ADP31592	Adp31592	Human	sec	236	18	58.1	1435	6	ABO00101	ABo00101	Novel	hum
164	18	58.1	1059	8	ADP31042	Adp31042	Human	sec	237	18	58.1	1435	6	ABU87112	ABu87112	Novel	hum
165	18	58.1	1066	8	ADP30563	Adp30563	Human	sec	238	18	58.1	1435	6	ABU91353	ABu91353	Novel	hum
166	18	58.1	1076	6	ABU81145	ABu81145	Human	PRO	239	18	58.1	1435	6	ABU90446	ABu90446	Novel	hum
167	18	58.1	1076	6	ABU6845	ABu6845	Human	PRO	240	18	58.1	1435	6	ABU97037	ABu97037	Novel	hum
168	18	58.1	1082	8	ADP30934	Adp30934	Human	sec	241	18	58.1	1435	6	ABO05233	ABo05233	Novel	hum
169	18	58.1	1086	8	ADP31175	Adp31175	Human	sec	242	18	58.1	1464	8	ADP31040	Adp31040	Human	sec
170	18	58.1	1090	8	ADQ10187	Adq10187	Human	pol	243	18	58.1	1480	8	ADP30557	Adp30557	Human	sec

244	18	58.1	1518	8	ADP31532	Adp31532 Human sec	317	18	58.1	2622	8	ADP31663	Adp31663 Human sec
245	18	58.1	1550	8	ADP30567	Adp30567 Human sec	318	18	58.1	2709	8	ADP31526	Adp31526 Human sec
246	18	58.1	1560	8	ADP31631	Adp31631 Human sec	319	18	58.1	2790	8	ADP31232	Adp31232 Human sec
247	18	58.1	1584	8	ADP31405	Adp31405 Human sec	320	18	58.1	2828	8	ADP30938	Adp30938 Human sec
248	18	58.1	1588	5	ABB09437	Abb09437 H. infIue	321	18	58.1	2833	8	ADP31299	Adp31299 Human sec
249	18	58.1	1614	8	ADP31529	Adp31529 Human sec	322	18	58.1	2835	8	ADP30572	Adp30572 Human sec
250	18	58.1	1617	8	ADP30660	Adp30660 Human sec	323	18	58.1	2976	8	ADP30724	Adp30724 Human sec
251	18	58.1	1647	8	ADP31052	Adp31052 Human sec	324	18	58.1	2980	9	AEA36049	Aea36049 Maize Sta
252	18	58.1	1662	8	ADP31419	Adp31419 Human sec	325	18	58.1	3036	8	ADP31595	Adp31595 Human sec
253	18	58.1	1662	8	ADP31513	Adp31513 Human sec	326	18	58.1	3144	8	ADP31544	Adp31544 Human sec
254	18	58.1	1665	8	ADP31187	Adp31187 Human sec	327	18	58.1	3201	8	ADP31545	Adp31545 Human sec
255	18	58.1	1719	8	ADP31137	Adp31137 Human sec	328	18	58.1	3339	8	ADP31219	Adp31219 Human sec
256	18	58.1	1725	8	ADP30654	Adp30654 Human sec	329	18	58.1	3350	8	ADP31148	Adp31148 Human sec
257	18	58.1	1733	6	ABU88255	Abu88255 Novel hum	330	18	58.1	3411	8	ADP30667	Adp30667 Human sec
258	18	58.1	1743	6	ABU90134	Abu90134 Novel hum	331	18	58.1	3447	8	ADP31112	Adp31112 Human sec
259	18	58.1	1743	6	ABU96436	Abu96436 Novel hum	332	18	58.1	3465	8	ADP31234	Adp31234 Human sec
260	18	58.1	1743	6	ABU99045	Abu99045 Novel hum	333	18	58.1	3579	8	ADP31098	Adp31098 Human sec
261	18	58.1	1743	6	ABU98260	Abu98260 Novel hum	334	18	58.1	4440	6	ABU88256	Abu88256 Novel hum
262	18	58.1	1743	6	ABU91966	Abu91966 Novel hum	335	18	58.1	4440	6	ABU90135	Abu90135 Novel hum
263	18	58.1	1743	6	ABU85270	Abu85270 Novel hum	336	18	58.1	4440	6	ABU96437	Abu96437 Novel hum
264	18	58.1	1743	6	ABO00409	Abu000409 Novel hum	337	18	58.1	4440	6	ABU99046	Abu99046 Novel hum
265	18	58.1	1743	6	ABU88960	Abu88960 Novel hum	338	18	58.1	4440	6	ABU98261	Abu98261 Novel hum
266	18	58.1	1743	6	ABO06456	Abu06456 Novel hum	339	18	58.1	4440	6	ABU91967	Abu91967 Novel hum
267	18	58.1	1743	6	ABU95516	Abu95516 Novel hum	340	18	58.1	4440	6	ABO85271	Abu085271 Novel hum
268	18	58.1	1743	6	ABU95206	Abu95206 Novel hum	341	18	58.1	4440	6	ABO00410	Abu000410 Novel hum
269	18	58.1	1743	6	ABU90754	Abu90754 Novel hum	342	18	58.1	4440	6	ABU88961	Abu88961 Novel hum
270	18	58.1	1743	6	ABU93916	Abu93916 Novel hum	343	18	58.1	4440	6	ABO06457	Abu06457 Novel hum
271	18	58.1	1743	6	ABU86190	Abu86190 Novel hum	344	18	58.1	4440	6	ABU95517	Abu95517 Novel hum
272	18	58.1	1743	6	ABU82045	Abu82045 Novel hum	345	18	58.1	4440	6	ABU95207	Abu95207 Novel hum
273	18	58.1	1743	6	ABU07906	Abu07906 Novel hum	346	18	58.1	4440	6	ABU90755	Abu90755 Novel hum
274	18	58.1	1743	6	ABU94226	Abu94226 Novel hum	347	18	58.1	4440	6	ABU93917	Abu93917 Novel hum
275	18	58.1	1743	6	ABO00099	Abu000099 Novel hum	348	18	58.1	4440	6	ABU96191	Abu96191 Novel hum
276	18	58.1	1743	6	ABU91351	Abu91351 Novel hum	349	18	58.1	4440	6	ABU82046	Abu82046 Novel hum
277	18	58.1	1743	6	ABU90444	Abu90444 Novel hum	350	18	58.1	4440	6	ABU97907	Abu97907 Novel hum
278	18	58.1	1743	6	ABU97035	Abu97035 Novel hum	351	18	58.1	4440	6	ABU94227	Abu94227 Novel hum
279	18	58.1	1743	6	ABO05231	Abu05231 Novel hum	352	18	58.1	4440	6	ABO00100	Abu000100 Novel hum
280	18	58.1	1743	6	ABU05231	Abu05231 Novel hum	353	18	58.1	4440	6	ABU87111	Abu87111 Novel hum
281	18	58.1	1746	8	ADP30992	Adp30992 Human sec	354	18	58.1	4440	6	ABU91352	Abu91352 Novel hum
282	18	58.1	1776	9	ADY62715	Ady62715 Human alp	355	18	58.1	4440	6	ABU90445	Abu90445 Novel hum
283	18	58.1	1782	8	ADP31391	Adp31391 Human sec	356	18	58.1	4440	6	ABU97036	Abu97036 Novel hum
284	18	58.1	1782	8	ADP31270	Adp31270 Human sec	357	18	58.1	4440	6	ABO05232	Abu005232 Novel hum
285	18	58.1	1789	8	ADP30962	Adp30962 Human sec	358	18	58.1	4752	8	ADP30585	Adp30585 Human sec
286	18	58.1	1793	4	ABB60964	Abb60964 Drcosphil	359	18	58.1	4752	8	ADP30651	Adp30651 Human sec
287	18	58.1	1815	8	ADP31601	Adp31601 Human sec	360	18	58.1	5304	8	ADP30706	Adp30706 Human sec
288	18	58.1	1827	8	ADP31170	Adp31170 Human sec	361	18	58.1	5397	8	ADP31068	Adp31068 Human sec
289	18	58.1	1833	8	ADP30642	Adp30642 Human sec	362	18	58.1	5514	8	ADP31186	Adp31186 Human sec
290	18	58.1	1933	8	ADP30889	Adp30889 Human sec	363	18	58.1	5514	8	ADP31591	Adp31591 Human sec
291	18	58.1	1933	8	ADP30902	Adp30902 Human sec	364	18	58.1	6465	8	ADP30705	Adp30705 Human sec
292	18	58.1	1956	8	ADP31662	Adp31662 Human sec	365	18	58.1	7285	6	ABJ38280	Abj38280 pAMG21-RA
293	18	58.1	2001	8	ADP31644	Adp31644 Human sec	366	18	58.1	7339	6	AAO16358	Aao16358 Human tra
294	18	58.1	2020	8	ADP31056	Adp31056 Human sec	367	18	58.1	8976	8	ADP31425	Adp31425 Human sec
295	18	58.1	2058	8	ADP31630	Adp31630 Human sec	368	18	58.1	9195	8	ADP31494	Adp31494 Human sec
296	18	58.1	2065	4	ABB63705	Abb63705 Drcosphil	369	18	58.1	10944	8	ADP31311	Adp31311 Human sec
297	18	58.1	2088	8	ADP31178	Adp31178 Human sec	370	18	58.1	11328	4	ABBS6666	Abbs6666 Human SNP
298	18	58.1	2091	8	ADP31088	Adp31088 Human sec	371	17	54.8	13	4	ABBS6666	Abbs6666 Human SNP
299	18	58.1	2127	8	ADP31327	Adp31327 Human sec	372	17	54.8	23	4	AAK12409	Aak12409 Albunin f
300	18	58.1	2148	8	ADP30974	Adp30974 Human sec	373	17	54.8	23	7	ADP68071	Adp68071 Human the
301	18	58.1	2187	8	ADP30882	Adp30882 Human sec	374	17	54.8	23	9	AEA39569	Aea39569 Nematode
302	18	58.1	22304	8	ADP31252	Adp31252 Human sec	375	17	54.8	23	9	AEA39570	Aea39570 Nematode
303	18	58.1	2347	8	ADP31394	Adp31394 Human sec	376	17	54.8	23	9	AEA39571	Aea39571 Nematode
304	18	58.1	2349	8	ADP30959	Adp30959 Human sec	377	17	54.8	24	7	ADD95109	Add95109 PCR prime
305	18	58.1	2391	8	ADP31366	Adp31366 Human sec	378	17	54.8	24	8	ADU59677	Adu59677 GBP-4 c10
306	18	58.1	2418	8	ADP31105	Adp31105 Human sec	379	17	54.8	25	2	AAV21389	Aav21389 Human HUP
307	18	58.1	2454	8	ADP30469	Adp30469 Human sec	380	17	54.8	28	4	AAV33045	Aam33045 peptide #
308	18	58.1	2484	8	ADP66690	Adp66690 Human sec	381	17	54.8	28	4	AAV72815	Aam72815 Human bon
309	18	58.1	2508	6	ADA5721	Ada5721 C. elegan	382	17	54.8	28	4	AAV60195	Aam60195 Human bira
310	18	58.1	2535	8	ADP31146	Adp31146 Human sec	383	17	54.8	28	4	ABG54519	Abg54519 Human liv
311	18	58.1	2542	8	ADP31594	Adp31594 Human sec	384	17	54.8	28	5	ABG42643	Abg42643 Human pep
312	18	58.1	2544	6	ADA15717	Ada15717 C. elegan	385	17	54.8	29	8	ADP30951	Adp30951 Human sec
313	18	58.1	2547	8	ADP31665	Adp31665 Human sec	386	17	54.8	31	3	AAK55811	Aak55811 Arabidops
314	18	58.1	2601	6	ADA15723	Ada15723 C. elegan	387	17	54.8	33	9	AEA34230	Aea34230 Opium pop
315	18	58.1	2616	8	ADP31253	Adp31253 Human sec	388	17	54.8	35	3	AAV99909	Aav99909 Peptide e
316	18	58.1	2616	9	AEA9675	Aeb9675 N. mening	389	17	54.8	45	6	ABR98384	Abt98384 Tumour ce

390	17	54.8	45	7	ADC64730	Adc64730	MCF-7	brc	463	17	54.8	157	7	ABO81786	ABO81786	Pseudomon
391	17	54.8	48	2	AAV13202	Aay13202	Human	sec	464	17	54.8	158	7	ABO75923	ABO75923	Pseudomon
392	17	54.8	48	9	AEA34232	Aea34232	Opium pop		465	17	54.8	161	4	AAH21861	AAH21861	Peptide #
393	17	54.8	49	8	ABO56643	Abos6643	Human	gen	466	17	54.8	161	4	ABH44230	ABH44230	Peptide #
394	17	54.8	51	4	AAU14915	Aau14915	Novel	bon	467	17	54.8	161	4	AAH38186	AAH38186	Peptide #
395	17	54.8	51	4	AAU50747	Aau50747	Proprionib		468	17	54.8	161	4	ABH27105	ABH27105	Protein #
396	17	54.8	51	6	ABM47266	Abm47266	Proprionib		469	17	54.8	161	4	AAH7967	AAH7967	Human bon
397	17	54.8	52	4	AAUS1165	Aaus1165	Proprionib		470	17	54.8	161	4	AAH65265	AAH65265	Human bra
398	17	54.8	52	6	ABM47684	Abm47684	Proprionib		471	17	54.8	161	4	ABG59605	ABG59605	Human liv
399	17	54.8	54	4	AAU44785	Aau44785	Proprionib		472	17	54.8	164	5	ABG46980	ABG46980	Human pep
400	17	54.8	54	6	ABM41304	Abm41304	Proprionib		473	17	54.8	164	8	ADH8412	ADH8412	Plant ful
401	17	54.8	58	5	ABP07833	Abp07833	Human ORP		474	17	54.8	164	8	ADH75562	ADH75562	Plant ful
402	17	54.8	61	4	AAH19250	Aam19250	Peptide #		475	17	54.8	165	8	ADP30587	ADP30587	Human sec
403	17	54.8	61	4	ABH38436	Abh38436	Peptide #		476	17	54.8	165	8	ADP31041	ADP31041	Human sec
404	17	54.8	61	4	AAH31941	Aam31941	Peptide #		477	17	54.8	167	9	AEA20070	AEA20070	Novel hum
405	17	54.8	61	4	ABH23648	Abh23648	Protein #		478	17	54.8	168	6	ABH39095	ABH39095	Molecule
406	17	54.8	61	4	AAH71647	Aam71647	Human bon		479	17	54.8	168	7	ADH04991	ADH04991	Human pro
407	17	54.8	61	4	AAH59112	Aam59112	Human bra		480	17	54.8	168	8	ADP31099	ADP31099	Human sec
408	17	54.8	61	4	ABG53331	Abg53331	Human liv		481	17	54.8	168	8	ADP30797	ADP30797	Human sec
409	17	54.8	61	5	ABG41461	Abg41461	Human pep		482	17	54.8	169	7	ABH83767	ABH83767	Pseudomon
410	17	54.8	68	4	AAH24038	Aam24038	Human EST		483	17	54.8	173	4	AAH64434	AAH64434	Proprionib
411	17	54.8	68	4	ADP31694	Adp31694	Human sec		484	17	54.8	173	6	ABH60953	ABH60953	Proprionib
412	17	54.8	74	2	AAH94646	Aaw94646	TNF-R ext		485	17	54.8	173	6	ABH41796	ABH41796	Human DIT
413	17	54.8	74	4	AAH69198	Aab69198	Human TNF		486	17	54.8	173	7	ABO83965	ABO83965	Pseudomon
414	17	54.8	74	7	ADH86259	Adh86259	Enterococ		487	17	54.8	173	9	AEA20072	AEA20072	Novel hum
415	17	54.8	74	7	ABO79221	AbO79221	Pseudomon		488	17	54.8	175	4	ABH68003	ABH68003	Drosophill
416	17	54.8	78	7	ADH18879	Adh18879	Human cel		489	17	54.8	175	7	ABO80411	ABO80411	Pseudomon
417	17	54.8	79	6	ABH18879	Adh18879	Human cel		490	17	54.8	177	6	ABH75865	ABH75865	Human sec
418	17	54.8	79	7	ADH35729	Adh35729	Human hep		491	17	54.8	177	8	ADP30755	ADP30755	Human sec
419	17	54.8	81	7	ABO79675	AbO79675	Pseudomon		492	17	54.8	180	8	ADP30820	ADP30820	Human sec
420	17	54.8	84	3	AAH99930	Aay99930	Peptide e		493	17	54.8	180	8	ADP30825	ADP30825	Human sec
421	17	54.8	86	7	ADH35728	Adh35728	Human hep		494	17	54.8	180	8	ADP31496	ADP31496	Human sec
422	17	54.8	86	7	ADH35728	Adh35728	Human hep		495	17	54.8	180	8	ADP30828	ADP30828	Human sec
423	17	54.8	90	8	ADP31655	Adp31655	Human sec		496	17	54.8	180	8	ADP30821	ADP30821	Human sec
424	17	54.8	92	4	AAH93635	Aau93635	Proprionib		497	17	54.8	180	8	ADP30826	ADP30826	Human sec
425	17	54.8	92	6	ABM6154	Abm6154	Proprionib		498	17	54.8	181	7	ABO73220	ABO73220	Pseudomon
426	17	54.8	93	4	AAU47955	Aau47955	Proprionib		499	17	54.8	181	7	ABO73421	ABO73421	Pseudomon
427	17	54.8	93	6	ABM44474	Abm44474	Proprionib		500	17	54.8	184	7	ABH77341	ABH77341	Pseudomon
428	17	54.8	93	8	ADP30859	Adp30859	Human sec		501	17	54.8	184	7	ABH88295	ABH88295	Rice abio
429	17	54.8	96	3	ADG27198	Aag27198	Zea maye		502	17	54.8	184	9	ABH20071	ABH20071	Novel hum
430	17	54.8	101	4	ABG12358	Abg12358	Novel hum		503	17	54.8	185	8	ADP31109	ADP31109	Human sec
431	17	54.8	103	7	ABO80232	AbO80232	Pseudomon		504	17	54.8	189	8	ADP30641	ADP30641	Human sec
432	17	54.8	107	4	AAU67241	Aau67241	Proprionib		505	17	54.8	191	4	AAU34753	AAU34753	Proprionib
433	17	54.8	107	6	ABM63760	Abm63760	Proprionib		506	17	54.8	191	6	ABH40272	ABH40272	Proprionib
434	17	54.8	108	3	AAH18584	Aag18584	Arabidops		507	17	54.8	195	8	ADP30590	ADP30590	Human sec
435	17	54.8	108	3	AAH15332	Aag15332	Arabidops		508	17	54.8	195	8	ADP30636	ADP30636	Human sec
436	17	54.8	109	4	ABH11092	Abh11092	Human sec		509	17	54.8	195	8	ADP30837	ADP30837	Human sec
437	17	54.8	109	8	ADH78125	Adh78125	Plant ful		510	17	54.8	203	7	ABO75097	ABO75097	Pseudomon
438	17	54.8	112	4	AAH45902	Aau45902	Proprionib		511	17	54.8	204	7	ABH90226	ABH90226	Rice abio
439	17	54.8	112	6	ABH42421	Abm42421	Proprionib		512	17	54.8	204	8	ADP31421	ADP31421	Human sec
440	17	54.8	114	5	ABH06472	Abp06472	Human ORP		513	17	54.8	204	8	ADP31422	ADP31422	Human sec
441	17	54.8	115	3	AAH54713	Aag54713	Arabidops		514	17	54.8	204	8	ADP31424	ADP31424	Human sec
442	17	54.8	117	3	AAH38749	Aag38749	Arabidops		515	17	54.8	204	8	ADP30545	ADP30545	Human sec
443	17	54.8	117	7	ADH60202	Adh60202	Secrete		516	17	54.8	204	8	ADP31420	ADP31420	Human sec
444	17	54.8	123	8	ADP31326	Adp31326	Human sec		517	17	54.8	207	7	ADH03987	ADH03987	Human pro
445	17	54.8	128	4	AAH50194	Aau50194	Proprionib		518	17	54.8	207	7	ABO71507	ABO71507	Pseudomon
446	17	54.8	128	6	ABH46713	Abm46713	Proprionib		519	17	54.8	208	7	ABO76863	ABO76863	Pseudomon
447	17	54.8	133	4	ADP31069	Adp31069	Human sec		520	17	54.8	211	6	ABH41776	ABH41776	Human DIT
448	17	54.8	134	8	ABG23610	Abg23610	Novel hum		521	17	54.8	211	7	ADH31155	ADH31155	Human dia
449	17	54.8	134	8	ADH09969	Adh09969	Human pro		522	17	54.8	213	8	ADP31498	ADP31498	Human sec
450	17	54.8	135	8	ADP30703	Adp30703	Human sec		523	17	54.8	213	8	ADP31499	ADP31499	Human sec
451	17	54.8	136	2	AAH92924	Aay92924	Amino aci		524	17	54.8	214	8	ADH06889	ADH06889	Plant ful
452	17	54.8	138	7	ABO75291	AbO75291	Pseudomon		525	17	54.8	219	8	ADP31171	ADP31171	Human sec
453	17	54.8	138	8	ADP31231	Adp31231	Human sec		526	17	54.8	222	8	ADP31379	ADP31379	Human sec
454	17	54.8	140	3	AAH83933	Aay83933	Human sec		527	17	54.8	225	7	ADH1861	ADH1861	Transcrip
455	17	54.8	140	4	ABH13377	Abh13377	Human thr		528	17	54.8	230	8	ABO84376	ABO84376	Pseudomon
456	17	54.8	140	7	ADH33298	Adh33298	Human nov		529	17	54.8	234	8	ABO58717	ABO58717	Human gen
457	17	54.8	140	7	ABO77148	AbO77148	Pseudomon		530	17	54.8	237	8	ADH22944	ADH22944	Plant ful
458	17	54.8	151	7	ABO83144	AbO83144	Pseudomon		531	17	54.8	239	2	ADH17135	ADH17135	Eucalyptu
459	17	54.8	152	7	ABO71621	AbO71621	Pseudomon		532	17	54.8	240	2	AAH06827	AAH06827	Thrombomo
460	17	54.8	154	7	ABO76866	AbO76866	Pseudomon		533	17	54.8	240	8	ADP31579	ADP31579	Human sec
461	17	54.8	154	7	ABO67900	AbO67900	Pseudomon		534	17	54.8	240	8	ADH07292	ADH07292	Lettuce p
462	17	54.8	154	7	ABO74421	AbO74421	Pseudomon		535	17	54.8	240	8	ADH77706	ADH77706	Plant ful

536	17	54.8	242	2	AAR06828	Aar06828 Thrombomo	609	17	54.8	361	4	ABB11492	Abb11492 Human neu
537	17	54.8	243	8	ADP30983	Adp30983 Human sec	610	17	54.8	366	8	ADP31106	Adp31106 Human sec
538	17	54.8	243	8	ADP31343	Adp31343 Human sec	611	17	54.8	366	8	ADP31670	Adp31670 Human sec
539	17	54.8	243	9	ADM17137	Adm17137 Eucalyptu	612	17	54.8	367	8	ADP30950	Adp30950 Human sec
540	17	54.8	249	2	AAR06826	Aar06826 Thrombomo	613	17	54.8	373	5	ABB47351	Abb47351 Listeria
541	17	54.8	249	2	ABO83180	AbO83180 Pseudomon	614	17	54.8	381	8	ADP30655	Adp30655 Human sec
542	17	54.8	252	8	ADP31485	Adp31485 Human sec	615	17	54.8	382	8	ADP31216	Adp31216 Human sec
543	17	54.8	253	2	AAR06825	Aar06825 Thrombomo	616	17	54.8	384	8	ADP30656	Adp30656 Human sec
544	17	54.8	253	2	AAR04241	Aar04241 Thrombin-	617	17	54.8	388	6	ABU39484	Abu39484 Protein e
545	17	54.8	253	2	AAR8514	Aar8514 Thrombomo	618	17	54.8	390	8	ADP31584	Adp31584 Human sec
546	17	54.8	253	4	ABO8326	AbO8326 Human chr	619	17	54.8	393	8	ADP31345	Adp31345 Human sec
547	17	54.8	254	3	AAy83938	Aay83938 Human chr	620	17	54.8	395	8	ADP30909	Adp30909 Human sec
548	17	54.8	254	3	AAy83937	Aay83937 Human chr	621	17	54.8	399	7	ABO76549	AbO76549 Pseudomon
549	17	54.8	258	8	ADP30597	Adp30597 Human sec	622	17	54.8	399	8	ADP31313	Adp31313 Human sec
550	17	54.8	262	9	AEBA3167	Aeba3167 Zea maye	623	17	54.8	402	7	ABO82656	AbO82656 Pseudomon
551	17	54.8	264	8	ADP30788	Adp30788 Human sec	624	17	54.8	402	8	ADP31414	Adp31414 Human sec
552	17	54.8	264	8	ADP31412	Adp31412 Human sec	625	17	54.8	406	7	ABO69985	AbO69985 Pseudomon
553	17	54.8	264	8	ADP31527	Adp31527 Human sec	626	17	54.8	411	8	ADP31104	Adp31104 Human sec
554	17	54.8	267	8	ADP30822	Adp30822 Human sec	627	17	54.8	411	8	ABM84979	Abm84979 Human dia
555	17	54.8	270	8	ADP30500	Adp30500 Human sec	628	17	54.8	412	4	ABUS3134	Abu3134 Human tra
556	17	54.8	274	8	ADP31236	Adp31236 Human sec	629	17	54.8	414	8	ADP31477	Adp31477 Human sec
557	17	54.8	274	8	ADY22941	Ady22941 Plant ful	630	17	54.8	417	8	ADP31432	Adp31432 Human sec
558	17	54.8	275	2	AAR06832	Aar06832 Thrombomo	631	17	54.8	417	9	ADM17788	Adm17788 Pirus rad
559	17	54.8	275	2	AAR94609	Aar94609 Human rec	632	17	54.8	420	8	ADP31349	Adp31349 Human sec
560	17	54.8	276	7	ABO69382	AbO69382 Pseudomon	633	17	54.8	421	7	ABO75404	AbO75404 Pseudomon
561	17	54.8	277	2	AAR06834	Aar06834 Thrombomo	634	17	54.8	421	8	ADP31159	Adp31159 Human sec
562	17	54.8	278	4	AAU45578	Aau45578 Propionib	635	17	54.8	421	8	ADS23895	AdS23895 Bacterial
563	17	54.8	278	6	ABM42097	Abm42097 Propionib	636	17	54.8	423	8	ADP30819	Adp30819 Human sec
564	17	54.8	278	8	ADP31314	Adp31314 Human sec	637	17	54.8	423	8	ADP31323	Adp31323 Human sec
565	17	54.8	279	8	ADP31107	Adp31107 Human sec	638	17	54.8	429	8	ADP31100	Adp31100 Human sec
566	17	54.8	279	8	ADP31024	Adp31024 Human sec	639	17	54.8	430	7	ABO69654	AbO69654 Pseudomon
567	17	54.8	279	8	ADP31489	Adp31489 Human sec	640	17	54.8	435	8	ADP31102	Adp31102 Human sec
568	17	54.8	288	8	ADP31463	Adp31463 Human sec	641	17	54.8	437	8	ADP31403	Adp31403 Human sec
569	17	54.8	289	8	ADT58146	Adt58146 Plant pol	642	17	54.8	440	6	ADA48432	Ada48432 Rice prot
570	17	54.8	290	7	ABO70728	AbO70728 Pseudomon	643	17	54.8	448	4	ABG20341	Abg20341 Novel hum
571	17	54.8	291	7	ABO68749	AbO68749 Pseudomon	644	17	54.8	450	8	ADP31085	Adp31085 Human sec
572	17	54.8	293	5	ABB92518	Abb92518 Herbicida	645	17	54.8	453	8	ADP31316	Adp31316 Human sec
573	17	54.8	293	4	ADN74129	Adn74129 Thale cre	646	17	54.8	456	8	ADP31636	Adp31636 Human sec
574	17	54.8	294	4	AAU67478	Aau67478 Propionib	647	17	54.8	461	2	AAR45335	Aar45335 Thrombomo
575	17	54.8	294	6	ABM63997	Abm63997 Propionib	648	17	54.8	461	8	ADP31634	Adp31634 Human sec
576	17	54.8	294	8	ADP31473	Adp31473 Human sec	649	17	54.8	462	2	AAR45347	Aar45347 Thrombomo
577	17	54.8	297	7	ABO83844	AbO83844 Pseudomon	650	17	54.8	462	2	AAR45349	Aar45349 Thrombomo
578	17	54.8	299	6	ABU19784	Abu19784 Protein e	651	17	54.8	462	2	AAR45337	Aar45337 Thrombomo
579	17	54.8	300	8	ADP30775	Adp30775 Human sec	652	17	54.8	462	2	AAR45339	Aar45339 Thrombomo
580	17	54.8	309	8	ADP30862	Adp30862 Human sec	653	17	54.8	462	2	AAR45343	Aar45343 Thrombomo
581	17	54.8	309	8	ADP30873	Adp30873 Human sec	654	17	54.8	462	2	AAR45355	Aar45355 Thrombomo
582	17	54.8	310	9	AEA39559	Aea39559 Nematode	655	17	54.8	462	2	AAR45341	Aar45341 Thrombomo
583	17	54.8	312	8	ADP30476	Adp30476 Human sec	656	17	54.8	462	2	AAR45342	Aar45342 Thrombomo
584	17	54.8	315	7	ADD30795	Add30795 Plant yle	657	17	54.8	462	2	AAR45336	Aar45336 Thrombomo
585	17	54.8	316	8	ADI41783	Adi41783 Plant tra	658	17	54.8	462	2	AAR45348	Aar45348 Thrombomo
586	17	54.8	316	8	ADX94660	Adx94660 Plant ful	659	17	54.8	462	2	AAR45350	Aar45350 Thrombomo
587	17	54.8	318	8	ADP31135	Adp31135 Human sec	660	17	54.8	462	2	AAR45345	Aar45345 Thrombomo
588	17	54.8	320	8	ADP31607	Adp31607 Human sec	661	17	54.8	462	2	AAR45354	Aar45354 Thrombomo
589	17	54.8	320	8	ADP31649	Adp31649 Human sec	662	17	54.8	462	2	AAR45338	Aar45338 Thrombomo
590	17	54.8	330	8	ADP31050	Adp31050 Human sec	663	17	54.8	462	2	AAR45353	Aar45353 Thrombomo
591	17	54.8	331	1	AAEP82847	Aaep82847 Sequence	664	17	54.8	462	2	AAR45346	Aar45346 Thrombomo
592	17	54.8	333	9	ADP31442	Adp31442 Human sec	665	17	54.8	463	2	AAR45344	Aar45344 Thrombomo
593	17	54.8	334	9	ABM96161	Abm96161 M. xanthu	666	17	54.8	463	2	AAR45340	Aar45340 Thrombomo
594	17	54.8	338	8	ADL06030	Adl06030 M. catarrh	667	17	54.8	466	5	ABB93873	Abb93873 Herbicida
595	17	54.8	339	8	ADP30892	Adp30892 Human sec	668	17	54.8	471	8	ADP31567	Adp31567 Human sec
596	17	54.8	339	8	ADP30702	Adp30702 Human sec	669	17	54.8	472	8	ADP31222	Adp31222 Human sec
597	17	54.8	340	6	ABM66183	Abm66183 Propionib	670	17	54.8	475	2	AAR22032	Aar22032 Truncated
598	17	54.8	344	6	ABO01324	AbO01324 Human pro	671	17	54.8	476	2	AAR78725	Aar78725 Mature th
599	17	54.8	344	7	ADM26628	Adm26628 Hyperther	672	17	54.8	476	2	AAR86377	Aar86377 Modified
600	17	54.8	344	8	ADN96020	Adn96020 Human NOV	673	17	54.8	478	2	AAR86376	Aar86376 Modified
601	17	54.8	345	8	ADP31016	Adp31016 Human sec	674	17	54.8	478	8	ADP31007	Adp31007 Human sec
602	17	54.8	345	8	ADP31683	Adp31683 Human sec	675	17	54.8	480	2	AAR22013	Aar22013 Truncated
603	17	54.8	349	6	ABM64594	Abm64594 Propionib	676	17	54.8	480	8	ADP31484	Adp31484 Human sec
604	17	54.8	350	8	ADP317315	Adp317315 Plant ful	677	17	54.8	484	4	ABB63251	Abb63251 Drosophil
605	17	54.8	353	8	ADP31558	Adp31558 Human sec	678	17	54.8	484	6	AAR30102	Aar30102 Drosophil
606	17	54.8	356	4	ABG04360	Abg04360 Novel hum	679	17	54.8	484	9	ABE53795	Aeb53795 Drosophil
607	17	54.8	357	8	ADP30505	Adp30505 Human sec	680	17	54.8	486	2	AAR13877	Aar13877 Thrombin-
608	17	54.8	360	8	ADP31439	Adp31439 Human sec	681	17	54.8	490	5	AAE17531	Aae17531 Human thr

682	17	54.8	490	5	AAE23032	Human chr	755	17	54.8	564	9	ADP31194	ADP31194	Human sec
683	17	54.8	491	2	AAE24400	Recomblina	756	17	54.8	564	9	ADY70520	ADY70520	Human bec
684	17	54.8	494	2	AAE24400	Recomblina	757	17	54.8	567	9	ADP31395	ADP31395	Human bec
685	17	54.8	494	2	AAE24400	Recomblina	758	17	54.8	567	9	ADP31395	ADP31395	Human bec
686	17	54.8	494	2	AAE24400	Recomblina	759	17	54.8	568	1	AAE24400	AAE24400	Human chr
687	17	54.8	497	2	AAE24400	Recomblina	760	17	54.8	575	2	AAE24400	AAE24400	Human chr
688	17	54.8	497	2	AAE24400	Recomblina	761	17	54.8	575	2	AAE24400	AAE24400	Human chr
689	17	54.8	498	2	AAE24400	Recomblina	762	17	54.8	575	2	AAE24400	AAE24400	Human chr
690	17	54.8	498	2	AAE24400	Recomblina	763	17	54.8	575	2	AAE24400	AAE24400	Human chr
691	17	54.8	498	2	AAE24400	Recomblina	764	17	54.8	575	2	AAE24400	AAE24400	Human chr
692	17	54.8	500	3	AAE24400	Recomblina	765	17	54.8	575	2	AAE24400	AAE24400	Human chr
693	17	54.8	502	4	AAE24400	Recomblina	766	17	54.8	575	2	AAE24400	AAE24400	Human chr
694	17	54.8	510	8	AAE24400	Recomblina	767	17	54.8	575	2	AAE24400	AAE24400	Human chr
695	17	54.8	510	8	AAE24400	Recomblina	768	17	54.8	575	2	AAE24400	AAE24400	Human chr
696	17	54.8	515	2	AAE24400	Recomblina	769	17	54.8	575	2	AAE24400	AAE24400	Human chr
697	17	54.8	516	2	AAE24400	Recomblina	770	17	54.8	575	2	AAE24400	AAE24400	Human chr
698	17	54.8	516	2	AAE24400	Recomblina	771	17	54.8	575	2	AAE24400	AAE24400	Human chr
699	17	54.8	516	2	AAE24400	Recomblina	772	17	54.8	575	2	AAE24400	AAE24400	Human chr
700	17	54.8	516	2	AAE24400	Recomblina	773	17	54.8	575	2	AAE24400	AAE24400	Human chr
701	17	54.8	516	2	AAE24400	Recomblina	774	17	54.8	575	2	AAE24400	AAE24400	Human chr
702	17	54.8	516	2	AAE24400	Recomblina	775	17	54.8	575	2	AAE24400	AAE24400	Human chr
703	17	54.8	516	2	AAE24400	Recomblina	776	17	54.8	575	2	AAE24400	AAE24400	Human chr
704	17	54.8	516	2	AAE24400	Recomblina	777	17	54.8	575	2	AAE24400	AAE24400	Human chr
705	17	54.8	516	2	AAE24400	Recomblina	778	17	54.8	575	2	AAE24400	AAE24400	Human chr
706	17	54.8	516	2	AAE24400	Recomblina	779	17	54.8	575	2	AAE24400	AAE24400	Human chr
707	17	54.8	516	2	AAE24400	Recomblina	780	17	54.8	575	2	AAE24400	AAE24400	Human chr
708	17	54.8	516	2	AAE24400	Recomblina	781	17	54.8	575	2	AAE24400	AAE24400	Human chr
709	17	54.8	516	2	AAE24400	Recomblina	782	17	54.8	575	2	AAE24400	AAE24400	Human chr
710	17	54.8	516	2	AAE24400	Recomblina	783	17	54.8	575	2	AAE24400	AAE24400	Human chr
711	17	54.8	516	2	AAE24400	Recomblina	784	17	54.8	575	2	AAE24400	AAE24400	Human chr
712	17	54.8	516	2	AAE24400	Recomblina	785	17	54.8	575	2	AAE24400	AAE24400	Human chr



828	17	54.8	631	6	ABG72572	Abg72572	Human	thr	901	17	54.8	813	8	ADP31282	Adp31282	Human	sec
829	17	54.8	639	8	ADP31521	Adp31521	Human	sec	902	17	54.8	814	8	ADP30561	Adp30561	Human	sec
830	17	54.8	642	8	ADP31265	Adp31265	Human	sec	903	17	54.8	815	4	ABB68374	Abb68374	Drosophila	
831	17	54.8	654	8	ADP30666	Adp30666	Human	sec	904	17	54.8	821	8	ADP30679	Adp30679	Human	sec
832	17	54.8	658	8	ADP31226	Adp31226	Human	sec	905	17	54.8	821	8	ADP30680	Adp30680	Human	sec
833	17	54.8	666	8	ADP31547	Adp31547	Human	sec	906	17	54.8	828	8	ADP31569	Adp31569	Human	sec
834	17	54.8	666	8	ADP30867	Adp30867	Human	sec	907	17	54.8	831	8	ADP31033	Adp31033	Human	sec
835	17	54.8	669	8	ADP31142	Adp31142	Human	sec	908	17	54.8	831	8	ADP31333	Adp31333	Human	sec
836	17	54.8	669	8	ADP31598	Adp31598	Human	sec	909	17	54.8	831	8	ADP31123	Adp31123	Human	sec
837	17	54.8	669	8	ADP31493	Adp31493	Human	sec	910	17	54.8	831	8	ADP31179	Adp31179	Human	sec
838	17	54.8	670	8	ADP31001	Adp31001	Human	sec	911	17	54.8	843	8	ADP30663	Adp30663	Human	sec
839	17	54.8	675	8	ADP31438	Adp31438	Human	sec	912	17	54.8	852	8	ADP30969	Adp30969	Human	sec
840	17	54.8	677	8	ADP31585	Adp31585	Human	sec	913	17	54.8	852	8	ADP30664	Adp30664	Human	sec
841	17	54.8	681	8	ADN05602	Adn05602	Antiprosi		914	17	54.8	856	8	ADP30972	Adp30972	Human	sec
842	17	54.8	681	8	ADP31053	Adp31053	Human	sec	915	17	54.8	858	4	ABB60597	Abb60597	Drosophila	
843	17	54.8	687	7	ABO77441	AbO77441	Pseudomon		916	17	54.8	868	8	ADP30736	Adp30736	Human	sec
844	17	54.8	690	8	ADP31686	Adp31686	Human	sec	917	17	54.8	870	8	ADP30646	Adp30646	Human	sec
845	17	54.8	690	8	ADP30750	Adp30750	Human	sec	918	17	54.8	876	8	ADP31220	Adp31220	Human	sec
846	17	54.8	702	8	ADP31518	Adp31518	Human	sec	919	17	54.8	882	8	ADP30487	Adp30487	Human	sec
847	17	54.8	711	8	ADP31215	Adp31215	Human	sec	920	17	54.8	882	8	ADP31688	Adp31688	Human	sec
848	17	54.8	711	8	ADP31652	Adp31652	Human	sec	921	17	54.8	885	8	ADP31198	Adp31198	Human	sec
849	17	54.8	711	8	ADP31535	Adp31535	Human	sec	922	17	54.8	887	8	ADP30554	Adp30554	Human	sec
850	17	54.8	714	8	ADP31561	Adp31561	Human	sec	923	17	54.8	887	8	ADP30548	Adp30548	Human	sec
851	17	54.8	720	5	AAO14994	Aao14994	Laminin-5		924	17	54.8	891	8	ADP31230	Adp31230	Human	sec
852	17	54.8	720	9	ADM47628	Adm47628	Laminin-5		925	17	54.8	894	8	ADP30851	Adp30851	Human	sec
853	17	54.8	725	8	ADP31092	Adp31092	Human	sec	926	17	54.8	900	8	ADP31337	Adp31337	Human	sec
854	17	54.8	727	7	ABM68246	Abm68246	Rice	abio	927	17	54.8	908	6	ABU08492	Abu08492	Alpha-hel	
855	17	54.8	728	8	ADP30508	Adp30508	Human	sec	928	17	54.8	918	8	ADP31459	Adp31459	Human	sec
856	17	54.8	731	4	ABG29843	Abg29843	Novel hum		929	17	54.8	925	8	ADP30546	Adp30546	Human	sec
857	17	54.8	735	8	ADP31520	Adp31520	Human	sec	930	17	54.8	925	5	AAO14246	Aao14246	Human	sec
858	17	54.8	739	4	ABR70356	Abbr70356	Drosophila		931	17	54.8	928	6	ABU20097	Abu20097	Protein e	
859	17	54.8	739	8	ADP31196	Adp31196	Human	sec	932	17	54.8	933	8	ADP31140	Adp31140	Human	sec
860	17	54.8	745	8	ADP30982	Adp30982	Human	sec	933	17	54.8	936	8	ADP31568	Adp31568	Human	sec
861	17	54.8	753	8	ADP30988	Adp30988	Human	sec	934	17	54.8	939	8	ADP31542	Adp31542	Human	sec
862	17	54.8	755	8	ADP31559	Adp31559	Human	sec	935	17	54.8	939	8	ADP31541	Adp31541	Human	sec
863	17	54.8	757	8	ADP30925	Adp30925	Human	sec	936	17	54.8	947	8	ADP30937	Adp30937	Human	sec
864	17	54.8	759	8	ADP31141	Adp31141	Human	sec	937	17	54.8	950	8	ADP31167	Adp31167	Human	sec
865	17	54.8	762	8	ADP30891	Adp30891	Human	sec	938	17	54.8	957	8	ADP31528	Adp31528	Human	sec
866	17	54.8	765	8	ADP31149	Adp31149	Human	sec	939	17	54.8	981	8	ADP30547	Adp30547	Human	sec
867	17	54.8	768	2	AAK27684	Aak27684	Human	bel	940	17	54.8	992	8	ADP31057	Adp31057	Human	sec
868	17	54.8	769	5	ABG96351	Abg96351	Human	ova	941	17	54.8	996	8	ADP31538	Adp31538	Human	sec
869	17	54.8	769	6	ABU56723	Abu56723	Lung	canc	942	17	54.8	1002	8	ADG39639	Adg39639	Human	pan
870	17	54.8	769	7	ADNB80492	Adnb80492	Ovarian c		943	17	54.8	1008	8	ADP30721	Adp30721	Human	sec
871	17	54.8	769	7	ADN39527	Adn39527	Cancer/an		944	17	54.8	1010	8	ADP31296	Adp31296	Human	sec
872	17	54.8	769	7	ADN39162	Adn39162	Cancer/an		945	17	54.8	1017	8	ADP31268	Adp31268	Human	sec
873	17	54.8	769	7	ADN39507	Adn39507	Cancer/an		946	17	54.8	1030	8	ADP30913	Adp30913	Human	sec
874	17	54.8	769	8	ADNB6616	Adnb6616	Human	int	947	17	54.8	1038	8	ADP30860	Adp30860	Human	sec
875	17	54.8	771	8	ADP31244	Adp31244	Human	sec	948	17	54.8	1041	8	ADP30998	Adp30998	Human	sec
876	17	54.8	772	8	ADP30936	Adp30936	Human	sec	949	17	54.8	1048	8	ADP31642	Adp31642	Human	sec
877	17	54.8	774	8	ADP30506	Adp30506	Human	sec	950	17	54.8	1050	8	ADP31376	Adp31376	Human	sec
878	17	54.8	774	8	ADP31373	Adp31373	Human	sec	951	17	54.8	1053	8	ADP30886	Adp30886	Human	sec
879	17	54.8	774	8	ADP31225	Adp31225	Human	sec	952	17	54.8	1064	7	ABO68762	AbO68762	Pseudomon	
880	17	54.8	775	7	ABM65895	Abm65895	Rice	abio	953	17	54.8	1065	8	ADP31347	Adp31347	Human	sec
881	17	54.8	779	8	ADP30915	Adp30915	Human	sec	954	17	54.8	1065	8	ADP31482	Adp31482	Human	sec
882	17	54.8	779	8	ADP30897	Adp30897	Human	sec	955	17	54.8	1065	8	ADP31287	Adp31287	Human	sec
883	17	54.8	780	2	AAK06824	Aak06824	Thrombomo		956	17	54.8	1065	8	ADP310970	Adp310970	Human	sec
884	17	54.8	780	8	ADP30770	Adp30770	Human	sec	957	17	54.8	1070	8	ADP31447	Adp31447	Human	sec
885	17	54.8	782	8	ADP30901	Adp30901	Human	sec	958	17	54.8	1086	8	ADP31447	Adp31447	Human	sec
886	17	54.8	783	8	ADP31284	Adp31284	Human	sec	959	17	54.8	1099	8	ADP30574	Adp30574	Human	sec
887	17	54.8	783	8	ADP31436	Adp31436	Human	sec	960	17	54.8	1104	8	ADP31156	Adp31156	Human	sec
888	17	54.8	783	8	ADP31398	Adp31398	Human	sec	961	17	54.8	1109	8	ADP31452	Adp31452	Human	sec
889	17	54.8	791	4	ABBS58225	Abbs58225	Drosophila		962	17	54.8	1125	8	ADP30920	Adp30920	Human	sec
890	17	54.8	796	8	ADP32123	Adp32123	PRO polyP		963	17	54.8	1128	6	ADA15725	Ada15725	C. elegans	
891	17	54.8	799	2	AAW02194	Aaw02194	Human	int	964	17	54.8	1134	8	ADP30647	Adp30647	Human	sec
892	17	54.8	799	5	AAU76337	Aau76337	Human	ant	965	17	54.8	1134	8	ADP31537	Adp31537	Human	sec
893	17	54.8	799	7	ADN95792	Adn95792	Human	BBC	966	17	54.8	1134	8	ADP30924	Adp30924	Human	sec
894	17	54.8	799	7	ADP65252	Adp65252	Human	int	967	17	54.8	1140	8	ADP31128	Adp31128	Human	sec
895	17	54.8	799	8	ADL06537	Adl06537	Human	tum	968	17	54.8	1140	8	ADP31130	Adp31130	Human	sec
896	17	54.8	799	8	ADP31261	Adp31261	Human	sec	969	17	54.8	1141	8	ADP30677	Adp30677	Human	sec
897	17	54.8	799	8	ADQ19150	Adq19150	Human	sec	970	17	54.8	1142	8	ADP30929	Adp30929	Human	sec
898	17	54.8	799	9	AEA81278	Aea81278	Human	int	971	17	54.8	1147	8	ADP30965	Adp30965	Human	sec
899	17	54.8	804	8	ADP31635	Adp31635	Human	sec	972	17	54.8	1171	8	ADP30906	Adp30906	Human	sec
900	17	54.8	807	8	ADP31036	Adp31036	Human	sec	973	17	54.8	1171	8	ADP30980	Adp30980	Human	sec

```
974 17 54.8 1173 8 ADP31155 Human sec Adp31155 Human sec
975 17 54.8 1179 8 ADP30486 Human sec Adp30486 Human sec
976 17 54.8 1179 8 ADP30577 Human sec Adp30577 Human sec
977 17 54.8 1183 8 ADP30536 Human sec Adp30536 Human sec
978 17 54.8 1192 8 ADP31180 Human sec Adp31180 Human sec
979 17 54.8 1194 8 ADP30682 Human sec Adp30682 Human sec
980 17 54.8 1197 8 ADP31034 Human sec Adp31034 Human sec
981 17 54.8 1197 8 ADP31342 Human sec Adp31342 Human sec
982 17 54.8 1200 9 ADZ76047 Mouse pre Adz76047 Mouse pre
983 17 54.8 1222 8 ADP30501 Human sec Adp30501 Human sec
984 17 54.8 1227 8 ADP31602 Human sec Adp31602 Human sec
985 17 54.8 1239 8 ADP31297 Human sec Adp31297 Human sec
986 17 54.8 1248 8 ADP31346 Human sec Adp31346 Human sec
987 17 54.8 1252 8 ADP30678 Human sec Adp30678 Human sec
988 17 54.8 1269 8 ADP31382 Human sec Adp31382 Human sec
989 17 54.8 1269 8 ADP31500 Human sec Adp31500 Human sec
990 17 54.8 1269 8 ADP31381 Human sec Adp31381 Human sec
991 17 54.8 1269 8 ADU23456 Sulfolobu Adu23456 Sulfolobu
992 17 54.8 1282 8 ADP31328 Human sec Adp31328 Human sec
993 17 54.8 1289 8 ADP30675 Human sec Adp30675 Human sec
994 17 54.8 1300 6 ABU88254 Novel hum Abu88254 Novel hum
995 17 54.8 1300 6 ABU90133 Novel hum Abu90133 Novel hum
996 17 54.8 1300 6 ABU96435 Novel hum Abu96435 Novel hum
997 17 54.8 1300 6 ABU99044 Novel hum Abu99044 Novel hum
998 17 54.8 1300 6 ABU98259 Novel hum Abu98259 Novel hum
999 17 54.8 1300 6 ABU91965 Novel hum Abu91965 Novel hum
1000 17 54.8 1300 6 ABU85269 Novel hum Abu85269 Novel hum
```

## ALIGNMENTS

```
RESULT 1
ID ABB41147 standard; peptide; 59 AA.
XX ABB41147;
AC ABB41147;
XX
XX 04-FEB-2002 (first entry)
XX
DE Peptide #8653 encoded by human foetal liver single exon probe.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
XX Homo sapiens.
XX
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX
XX 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR,
XX
XX WPI; 2001-483447/52.
XX
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX
XX gene expression in human foetal liver.
XX
XX
XX Claim 27; SEQ ID NO 33782; 639pp + Sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
XX
XX human gene expression in a sample derived from human foetal liver. The
```

```
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at fcp.wipo.int/pub/published\_pct\_sequences
XX
XX Sequence 59 AA;
```

```
Query Match 58.1%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
Qy 4 CXXXXXXC 12
Db 13 CXXSSTSTC 21

RESULT 2
ID AAM34923 standard; protein; 59 AA.
XX AAM34923;
AC AAM34923;
XX
XX 17-OCT-2001 (first entry)
XX
XX Peptide #8960 encoded by probe for measuring placental gene expression.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX
XX genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR,
XX
XX WPI; 2001-488897/53.
XX
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX
XX gene expression in human placenta.
XX
XX
XX Claim 27; SEQ ID NO 35192; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SNP:
XX
XX see AAI31315-AA157546). The present sequence is a peptide encoded by one
XX
XX such probe. The probes are useful for producing a microarray for
XX
XX predicting, measuring and displaying gene expression in samples derived
XX
XX from human placenta. The probes are useful for antenatal diagnosis of
XX
XX human genetic disorders
XX
XX Sequence 59 AA;
SQ
Qy Query Match 58.1%; Score 18; DB 4; Length 59;
Best Local Similarity 22.2%; Pred. No. 2.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

Db 13 CSSSSTSTC 21

RESULT 3  
ID ABB25185 standard; protein; 59 AA.  
XX ABB25185;  
AC  
XX 23-JAN-2002 (first entry)  
XX  
XX Protein #7184 encoded by probe for measuring heart cell gene expression.  
DE  
XX  
XX Human: Gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W0200157274-A2.  
PN  
XX  
XX 09-AUG-2001.  
PD  
XX  
XX 30-JAN-2001; 2001WO-US000666.  
PF  
XX  
XX 04-FEB-2000; 2000US-0180312P.  
PR  
XX 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX WPI; 2001-488990/53.  
DR  
XX  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts.  
XX  
XX  
XX Claim 15; SEQ ID NO 26955; 530pp; English.  
PS  
XX  
XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ffp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 59 AA;  
SQ

Query Match 58.1%; Score 18; DB 4; Length 59;  
Best Local Similarity 22.2%; Pred. No. 2.8e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXC 12  
|  
Db 13 CSSSSTSTC 21

RESULT 4  
ID AAM74807 standard; protein; 59 AA.  
XX  
XX AAM74807;  
AC

XX  
DT 06-NOV-2001 (first entry)  
XX  
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 35113.  
DE  
XX Human: bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukemia; lymphoma; myeloma.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W0200157276-A2.  
PN  
XX  
XX 09-AUG-2001.  
PD  
XX  
XX 30-JAN-2001; 2001WO-US000666.  
PF  
XX  
XX 04-FEB-2000; 2000US-0180312P.  
PR  
XX 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX WPI; 2001-488990/53.  
DR  
XX  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human bone marrow.  
XX  
XX  
XX Example 4; SEQ ID NO 35113; 658pp + Sequence Listing; English.  
PS  
XX  
XX The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention  
XX  
XX  
XX Sequence 59 AA;  
SQ

Query Match 58.1%; Score 18; DB 4; Length 59;  
Best Local Similarity 22.2%; Pred. No. 2.8e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXC 12  
|  
Db 13 CSSSSTSTC 21

RESULT 5  
ID AAM62003 standard; protein; 59 AA.  
XX  
XX AAM62003;  
AC  
XX  
XX 05-NOV-2001 (first entry)  
DT  
XX  
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 34108.  
DE  
XX Human: brain expressed exon; gene expression analysis; probe; microarray;  
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W0200157275-A2.  
PN  
XX  
XX 09-AUG-2001.  
PD  
XX  
XX 30-JAN-2001; 2001WO-US000667.  
PF

XX 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
DR WPI; 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains.  
XX  
PS Example 4; SEQ ID NO 34108; 650bp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system,  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention  
XX  
SQ Sequence 59 AA;

Query Match 58.1%; Score 18; DB 4; Length 59;  
Best Local Similarity 22.2%; Pred. No. 2.8e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 4 CXXXXXXC 12  
Db 13 CXXXXXXC 21

RESULT 6  
ABG56589  
ID ABG56589 standard; peptide; 59 AA.  
XX  
AC ABG56589;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human liver peptide, SEQ ID No 35237.  
XX  
KW Human, liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
KW hypercholesterolaemia; coronary heart disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200157273-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000664.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488998/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human adult liver.  
XX  
PS Claim 27; SEQ ID NO 35237; 658bp; English.

CC The invention relates to a single exon nucleic acid probe (SNP) (1) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult liver.  
CC (1) may be used for predicting, measuring and displaying gene expression  
CC in samples derived from human adult liver. The genes identified may be  
CC involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
CC associated with coronary heart disease. ABG47348-ABG59930 represent human  
CC liver single exon encoded peptides of the invention. Note: The sequence  
CC information for this patent does not appear in the printed specification  
CC but was obtained in electronic format directly from WPIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

Query Match 58.1%; Score 18; DB 4; Length 59;  
Best Local Similarity 22.2%; Pred. No. 2.8e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 4 CXXXXXXC 12  
Db 13 CXXXXXXC 21

RESULT 7  
ABG44598  
ID ABG44598 standard; peptide; 59 AA.  
XX  
AC ABG44598;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human peptide encoded by genome-derived single exon probe SEQ ID 34263.  
XX  
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;  
KW chronic obstructive pulmonary disease; interstitial lung disease;  
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;  
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KW primary ciliary dyskinesia; pulmonary hypertension;  
KW hyaline membrane disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200186003-A2.  
XX  
PD 15-NOV-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000665.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples.  
 XX  
 PS Claim 27; SEQ ID NO 34263; 634bp; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of probes  
 CC; the novel set of probes which hybridize at high stringency to a nucleic  
 CC acid expressed in the human lung; measuring gene expression in a sample  
 CC derived from human lung, comprising (a) contacting the array with a  
 CC collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of the  
 CC array; identifying exons in a eukaryotic genome, comprising (a)  
 CC algorithmically predicting at least one exon from genomic sequences of  
 CC the eukaryote; and (b) detecting specific hybridization of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridization to a single exon  
 CC microarray having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene expression  
 CC analysis, and for identifying exons in a gene, particularly using human  
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hernanaky-  
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Kargener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a peptide/protein encoded by a single exon probe of  
 CC the invention. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at [ftp.wipo.int/pub/published\\_pcr\\_sequences](http://ftp.wipo.int/pub/published_pcr_sequences)  
 CC  
 SQ Sequence 59 AA;  
 XX  
 CC Query Match 58.1%; Score 18; DB 5; Length 59;  
 CC Best Local Similarity 22.2%; Pred. No. 2.8e+02;  
 CC Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 CXXXXXXXC 12  
 DB 13 CXXXXXXXC 21  
 XX  
 RESULT 8  
 AAB21190  
 ID AAB21190 standard; protein; 108 AA.  
 XX  
 AC AAB21190;  
 XX  
 DT 08-JAN-2001 (first entry)  
 XX  
 DE Exo14 partial protein.  
 XX  
 CC Exocytosis pathway protein; Exo protein; antiparkinsonian; antidiabetic;  
 CC antiasthmatic; antiallergic; antitubercular; neuroprotective; anticonvulsant;  
 CC vulnary; asthma; inflammation; allergy; Chediak-Higashi syndrome; CHS;  
 CC Alzheimer's disease; Parkinson's disease; Huntington's disease; diabetes;  
 CC digestion disorder; wound healing disorder; gene therapy.  
 CC  
 OS Mus sp.

XX  
 PN WO200043419-A2.  
 XX  
 PD 27-JUL-2000.  
 XX  
 XX 20-JAN-2000; 2000WO-US001431.  
 XX  
 PR 20-JAN-1999; 99US-0116534P.  
 PR 26-JAN-1999; 99US-0117274P.  
 PR 26-JAN-1999; 99US-0117308P.  
 PR 26-JAN-1999; 99US-0117309P.  
 PR 26-JAN-1999; 99US-0117312P.  
 PR 01-FEB-1999; 99US-0118177P.  
 PR 01-FEB-1999; 99US-0118178P.  
 PR 01-FEB-1999; 99US-0118179P.  
 PR 09-FEB-1999; 99US-0119286P.  
 PR 11-FEB-1999; 99US-0119759P.  
 PR 11-FEB-1999; 99US-0119998P.  
 PA (RIGEL-) RIGEL PHARM INC.  
 XX  
 PI Luo Y;  
 XX  
 DR WPI; 2000-482908/42.  
 DR N-PSDB; AAA89569.  
 XX  
 PT New nucleic acids encoding Exo proteins which are useful in the  
 PT diagnosis, treatment or prevention of exocytosis-mediated disorders such  
 PT as asthma, inflammation and allergies.  
 XX  
 PS Disclosure; Page 149-150; 305pp; English.  
 XX  
 CC The present sequence is a polypeptide which is associated with the  
 CC exocytosis pathway. cDNA molecules encoding proteins involved in  
 CC exocytosis have been isolated by yeast one-hybrid and two-hybrid  
 CC screening. Novel proteins, termed Exo proteins, have been identified that  
 CC interact with known exocytosis-associated proteins such as GS27, alpha  
 CC snap, unc18-1, vamps, snap-23, and the rab family of proteins. Exo  
 CC proteins and their agonists and antagonists are useful in the diagnosis,  
 CC treatment or prevention of exocytosis-mediated disorders such as asthma,  
 CC inflammation, allergies, Chediak-Higashi Syndrome (CHS), Alzheimer's  
 CC disease, Parkinson's disease, Huntington's disease, diabetes, digestion  
 CC disorders and wound healing disorders. The nucleic acids, antagonists or  
 CC agonists of Exo proteins are useful in gene therapy. The nucleic acids  
 CC are also useful for generating transgenic or knock-out animals which can  
 CC be used in the development and screening of therapeutically useful  
 CC reagents  
 CC  
 SQ Sequence 108 AA;  
 XX  
 CC Query Match 58.1%; Score 18; DB 3; Length 108;  
 CC Best Local Similarity 22.2%; Pred. No. 3.2e+02;  
 CC Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 CXXXXXXXC 12  
 DB 33 CXXXXXXXC 41  
 XX  
 RESULT 9  
 ABO75321  
 ID ABO75321 standard; protein; 134 AA.  
 XX  
 AC ABO75321;  
 XX  
 DT 29-JUL-2004 (first entry)  
 XX  
 CC Pseudomonas aeruginosa polypeptide #7496.  
 CC Pseudomonas aeruginosa infection; Pseudomonas aeruginosa infection; antibacterial.  
 CC Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
 CC  
 OS Pseudomonas aeruginosa.  
 XX

PN	US6551795-B1.
XX	
PD	22-APR-2003.
XX	
PF	18-FEB-1999; 99US-00252991.
XX	
PR	18-FEB-1998; 98US-0074788P.
PR	27-JUL-1998; 98US-0094190P.
PA	(GENO-) GENOME THERAPEUTICS CORP.
PI	Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX	
DR	WPI; 2003-615309/58.
DR	N-PsDB; ABD08892.
XX	
PT	Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, of PT useful as molecular targets for diagnostics, prophylaxis and treatment of PT pathological conditions resulting from bacterial infection.
PS	Disclosure; SEQ ID NO 24067; 455pp; English.
XX	
CC	The invention relates to Pseudomonas aeruginosa polypeptides and the
CC	polynucleotides encoding them. The sequences are useful in diagnosis and
CC	therapy of pathological conditions, as molecular targets for diagnostics,
CC	prophylaxis and treatment of pathological conditions resulting from a
CC	bacterial infection, for evaluating a compound, such as a polypeptide,
CC	for the ability to bind a P. aeruginosa nucleic acid, as components of
CC	effective antibacterial targets, as targets for antibacterial drugs,
CC	including anti-P. aeruginosa drugs, as templates for recombinant
CC	production of P. aeruginosa-derived peptides or polypeptides, as target
CC	components for diagnosis and/or treatment of P. aeruginosa-caused
CC	infection, and in detection of P. aeruginosa sequences or other sequences
CC	of Pseudomonas species using biochip technology. Sequences ABO67826-
CC	ABO84336 represent P. aeruginosa polypeptides of the invention. Note: The
CC	sequence data for this patent did not form part of the printed
CC	spectification but was obtained in electronic format from USPRO at
CC	seqdata.uspto.gov/sequence.html
XX	
SQ	Sequence 134 AA;
Query Match	58.1%; Score 18; DB 7; Length 134;
Best Local Similarity	22.2%; Pred. NO. 3.4e+02;
Matches	2; Conservative 0; Mismatches 7; Indels 0; Gaps 0.
CY	4 CXXXXXXXXC 12
Db	19 CSASTTSC 27
RESULT 10	
ADE72390	
ID	ADE72390 standard; protein; 139 AA.
XX	
CC	ADE72390;
XX	
DT	29-JAN-2004 (first entry)
XX	
DE	Human endometrial specific protein, SEQ ID NO 430.
XX	
KW	cyclostatic; vaccine; human; endometrial specific genes; KW endometrial specific protein; endometrial cancer.
OS	Homo sapiens.
XX	
PN	WO2003060081-A2.
XX	
PD	24-JUL-2003.
XX	
PF	23-DEC-2002; 2002WO-US041612.
XX	
PR	21-DEC-2001; 2001US-0342756P.
XX	

```

PA (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Liu C;
XX
XX WPI; 2003-577666/54.
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
XX
XX Claim 12; SEQ ID NO 430; 824pp; English.
XX
XX
XX The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present amino acid sequence represents a
XX human endometrial specific protein of the invention.
XX
XX
XX Sequence 139 AA;
XX
XX
XX Query Match 58.1%; Score 18; DB 7; Length 139;
XX Best Local Similarity 22.2%; Pred. NO. 3.4e+02;
XX Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0
XX
XX 4 CXXXXXXC 12
XX |
XX 60 CSTSSAASC 68
XX
XX
XX RESULT 11
XX ADE72391
XX ID ADE72391 standard; protein, 139 AA.
XX
XX AC ADE72391;
XX
XX DT 29-JAN-2004 (first entry)
XX
XX Human endometrial specific protein, SEQ ID NO 431.
XX
XX cyostatic; vaccine; human; endometrial specific genes;
XX endometrial specific protein; endometrial cancer.
XX
XX OS Homo sapiens.
XX
XX FN WO2003060081-A2.
XX
XX PD 24-JUL-2003.
XX
XX PF 23-DEC-2002; 2002MO-US041612.
XX
XX PR 21-DEC-2001; 2001US-0342756P.
XX
XX PA (DIAD-) DIADEXUS INC.
XX
XX PI Sun Y, Liu C;
XX
XX WPI; 2003-577666/54.
XX
XX Nucleic acid molecules and polypeptides useful for diagnosing and
XX treating endometrial cancer and non-cancerous disease states in
XX endometrial.
XX
XX Claim 12; SEQ ID NO 431; 824pp; English.
XX
XX
XX The invention comprises the amino acid and DNA sequences of human
XX endometrial specific genes and proteins. The DNA and protein sequences of
XX the invention are useful for diagnosing, imaging and treating a patient
XX with endometrial cancer. The present amino acid sequence represents a
XX human endometrial specific protein of the invention.
XX
XX
XX Sequence 139 AA;
XX
XX
XX Query Match 58.1%; Score 18; DB 7; Length 139;

```

Best Local Similarity 22.2%; Pred. No. 3.4e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12  
Db 60 CSTSSAASC 68

## RESULT 12

AAM93094  
ID AAM93094 standard; protein, 140 AA.

AC AAM93094;

DT 06-NOV-2001 (first entry)

DE Human digestive system antigen SEQ ID NO: 2443.

XX Human; digestive system antigen; gene therapy; cancer; appendicitis;  
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;  
XX digestive system disorder; Meckel's diverticulum.

OS Homo sapiens.

PN WO200155314-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US001324.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0226799P.

PR 22-AUG-2000; 2000US-0226819P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239335P.

PR 13-OCT-2000; 2000US-0239337P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-502630/55.  
XX N-PSDB; AAK88867.  
XX  
XX Polynucleotides encoding digestive system antigens, useful for  
PT diagnosing, treating, preventing and/or prognostizing disorders of the  
PT digestive system, particularly cancer and cancer metastases.  
XX  
XX Claim 11; SEQ ID NO 2443; 986bp; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic  
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a digestive system antigen of  
CC the invention  
XX  
XX  
XX Sequence 140 AA;  
SQ  
Query Match 58.1%; Score 18; DB 4; Length 140;  
Best Local Similarity 22.2%; Pred. No. 3.4e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXC 12  
DB 104 CTATSATAC 112  
RESULT 13  
ID AAU20102 standard; protein; 140 AA.  
XX  
XX AAU20102;  
AC  
XX  
XX  
DT 06-DEC-2001 (first entry)  
XX  
DE Human liver associated polypeptide #133.  
XX  
XX Liver associated protein; human; mouse; rabbit; goat; horse; cat; dog;  
XX chicken; sheep; immunosuppressive; antihairtic; vasotropic;  
XX antihemetic; antiproliferative; cytostatic; cardiant; neuroprotective;  
XX cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
XX ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;  
XX hyperproliferative disorder; breast; liver; cardiovascular disorder;  
XX cerebrovascular disorder; nervous system disorder; bacterial infection;  
XX fungal infection; viral infection; ocular disorder; endocrine disorder;  
XX gastrointestinal disorder; renal disorder; respiratory disorder;  
XX wound healing; skin aging; organ transplantation; tissue regeneration;  
XX

KW anti-infertility.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200155355-A1.  
PN  
XX  
XX 02-AUG-2001.  
ED  
XX  
XX 17-JAN-2001; 2001WO-US001351.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0196123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229309P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.



PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241888P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251888P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251899P.  
 PR 08-DEC-2000; 2000US-0251900P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Barash SC, Ruben SM;  
 PI WPI; 2001-45728/49.  
 DR N-PSDB; AAS31813.  
 XX  
 PT Isolated nucleic acid molecule encoding a human liver related protein is  
 PT used in preventing, treating or ameliorating disorders of the liver  
 PT particularly cancer of the liver.  
 PS Claim 11; SEQ ID NO 289; 526pp; English.  
 XX  
 CC Sequences AAU9970-AAU2015 represent the liver associated polypeptides  
 CC of the invention. Liver associated polypeptides and their associated  
 CC polynucleotides are useful in the diagnosis, treatment and prevention of  
 CC various types of disorders in e.g. humans, mice, rabbits, goats, horses,  
 CC cats, dogs, chickens or sheep. A pathological condition can be determined  
 CC by detecting the presence or absence of a mutation in a liver associated  
 CC polynucleotide. The treatable disorders include autoimmune diseases such  
 CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms  
 CC of the breast or liver, cardiovascular disorders such as cardiac arrest,  
 CC cerebrovascular disorders such as cerebral ischaemia, nervous system  
 CC disorders such as Alzheimer's disease, infections caused by bacteria,  
 CC viruses and fungi, ocular disorders such as corneal infection, endocrine  
 CC disorders such as Crohn's disease, renal disorders such as  
 CC glomerulonephritis and respiratory disorders such as asthma and pleurisy.  
 CC The polypeptides can also be used to aid wound healing, to prevent skin  
 CC aging due to sunburn, to maintain organs before transplantation, to  
 CC regenerate tissues and in chemotaxis. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX

Query Match 58.1%; Score 18; DB 4; Length 140;  
 Best Local Similarity 22.2%; Pred. No. 3.4e+02;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
 Db 104 CTATSATAC 112

RESULT 14  
 ID ABP40963 standard; protein; 140 AA.  
 AC ABP40963;  
 XX  
 DT 24-JUL-2002 (first entry)  
 XX  
 DB Human liver antigen HFLV570, SEQ ID NO:289.  
 XX  
 KW Human; liver antigen; liver disorder; hepatic disorder; infection;  
 KW hepatitis; viral; parasitic; bacterial; fungal; inflammatory condition;  
 KW cirrhosis; granulomatous hepatitis; toxin damage; drug damage;  
 KW autoimmune disease; Wilson's disease; primary biliary cirrhosis;  
 KW neoplastic disorder; cancer; tumour; portal hypertension;  
 KW gastrointestinal disorder; hepatitis; drug screening; gene therapy;  
 KW chromosome mapping; forensic analysis; antibody preparation;  
 KW hepatotropic; cytotoxic; antiinflammatory; virucide; antibacterial;  
 KW fungicide; parasiticide; antidote; immunosuppressive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002042096-A1.  
 XX  
 PD 11-APR-2002.  
 XX  
 PF 17-JAN-2001; 2001US-00764887.

XX	PR	31-JAN-2000;	2000US-0179065P.
PR	04-FEB-2000;	2000US-0160628P.	
PR	28-JUN-2000;	2000US-0214886P.	
PR	07-JUL-2000;	2000US-0216647P.	
PR	07-JUL-2000;	2000US-0216880P.	
PR	11-JUL-2000;	2000US-0217487P.	
PR	11-JUL-2000;	2000US-0217496P.	
PR	14-JUL-2000;	2000US-0218290P.	
PR	26-JUL-2000;	2000US-022063P.	
PR	26-JUL-2000;	2000US-0220964P.	
PR	14-AUG-2000;	2000US-0224518P.	
PR	14-AUG-2000;	2000US-0224519P.	
PR	14-AUG-2000;	2000US-0225267P.	
PR	14-AUG-2000;	2000US-0225268P.	
PR	14-AUG-2000;	2000US-0225270P.	
PR	14-AUG-2000;	2000US-0225447P.	
PR	14-AUG-2000;	2000US-0225757P.	
PR	22-AUG-2000;	2000US-0225758P.	
PR	30-AUG-2000;	2000US-02286924P.	
PR	01-SEP-2000;	2000US-0229287P.	
PR	01-SEP-2000;	2000US-0229343P.	
PR	01-SEP-2000;	2000US-0229344P.	
PR	01-SEP-2000;	2000US-0229345P.	
PR	05-SEP-2000;	2000US-0229509P.	
PR	05-SEP-2000;	2000US-0229513P.	
PR	08-SEP-2000;	2000US-0231413P.	
PR	21-SEP-2000;	2000US-0234223P.	
PR	21-SEP-2000;	2000US-0234274P.	
PR	25-SEP-2000;	2000US-0234997P.	
PR	27-SEP-2000;	2000US-0235834P.	
PR	29-SEP-2000;	2000US-0236337P.	
PR	29-SEP-2000;	2000US-0236367P.	
PR	29-SEP-2000;	2000US-0236368P.	
PR	29-SEP-2000;	2000US-0236370P.	
PR	29-SEP-2000;	2000US-0236370P.	
PR	02-OCT-2000;	2000US-0236802P.	
PR	02-OCT-2000;	2000US-0237037P.	
PR	02-OCT-2000;	2000US-0237038P.	
PR	02-OCT-2000;	2000US-0237039P.	
PR	02-OCT-2000;	2000US-0237040P.	
PR	13-OCT-2000;	2000US-0239935P.	
PR	20-OCT-2000;	2000US-0240960P.	
PR	20-OCT-2000;	2000US-0241785P.	
PR	20-OCT-2000;	2000US-0241809P.	
PR	01-NOV-2000;	2000US-0244617P.	
PR	17-NOV-2000;	2000US-0249259P.	
PR	08-DEC-2000;	2000US-0251856P.	
PR	08-DEC-2000;	2000US-0251868P.	
PR	08-DEC-2000;	2000US-0251869P.	
XX	PA	(ROSE/) ROSEN C A.	
XX	PA	(RUBE/) RUBEN S M.	
XX	PA	(BARA/) BARASH S C.	
XX	PL	Rosen CA, Ruben SM, Barash SC;	
XX	DR	WPI; 2002-381944/41.	
XX	DR	N-PSDB; ABN90168.	
PT	PT	New nucleic acid encoding human liver antigens, useful for diagnosis,	
PT	PT	treatment and prevention of e.g. hepatitis and hepatic cancer, also	
PT	PT	related polypeptides and antibodies.	
XX	XX	Claim 11; SEQ ID NO 289; 181pp; English.	
PS	XX		
XX	XX		
CC	CC	The invention relates to 145 novel human liver antigens (ABP40831-	
CC	CC	ABP40975) and to cDNAs encoding them (ABN90036-ABN90180), and also	
CC	CC	encompasses polypeptides 90% identical and polynucleotides 95% identical	
CC	CC	to the sequences of the invention. The invention additionally relates to	
CC	CC	recombinant vectors and host cells comprising human liver antigen	
CC	CC	polynucleotides, antibodies against human liver antigens, and the use of	

XX	Sequence	140 AA;
XX	Query Match	58.1%; Score 18; DB 5; Length 140;
CC	Best Local Similarity	22.2%; Pred. No. 3.4e+02;
CC	Matches	2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY	4 CXXXXXXC 12	
Db	104 CTATSATAC 112	
RESULT 15		
ID	ADJ15081	
AD	ADJ15081 standard; protein: 140 AA.	
AC	ADJ15081;	
XX		
DT	20-MAY-2004 (first entry)	
XX		
DE	Human liver-related protein - SEQ ID 289.	
XX		
KW	liver; vitucide; fungicide; antibacterial; antiparasitic; hepatotropic;	
KW	antiflammatory; cytostatic; litholytic; antirheumatic; antiarthritic;	
KW	neuroprotective; antidiabetic; anticoagulant; thrombolytic;	
KW	antiartherosclerotic; cardiant; haemostatic; antiarrhythmic;	
KW	ophthalmological; antiarteriosclerotic; vasotropic; osteopathic;	
KW	nocotropic; antiparkinsonian; anticomulant; neuroleptic; vasotropic;	
KW	cytostatic; gynaecological; viral; fungal; bacterial;	
KW	parasitic infection; cirrhosis; Wilson's disease;	
KW	gastrointestinal disorder; pancreatic; gallbladder; immune; blood;	
KW	hyperproliferative; cardiovascular; respiratory; musculoskeletal system;	
KW	neurological; endocrine; reproductive system; developmental; inherited;	
XX	human.	
OS	Homo sapiens.	
XX		
PN	US2003077602-A1.	
XX		
PD	24-APR-2003.	
XX		
PF	14-FEB-2002; 2002US-00073961.	
XX		
PR	31-JAN-2000; 2000US-0179065P.	
PR	04-FEB-2000; 2000US-0180628P.	
PR	24-FEB-2000; 2000US-0184664P.	
PR	02-MAR-2000; 2000US-0186350P.	
PR	16-MAR-2000; 2000US-0189874P.	
PR	17-MAR-2000; 2000US-0190076P.	
PR	18-APR-2000; 2000US-0198123P.	
PR	19-MAY-2000; 2000US-0205515P.	



CC litholytic, antirheumatic, antiarthritic, neuroprotective, antidiabetic,  
CC anticoagulant, thrombolytic, antiarteriosclerotic, cardiant, haemostatic,  
CC antiarrhythmic, ophthalmological, antiarteriosclerotic, vasotropic,  
CC osteopathic, nootropic, antiparkinsonian, anticonvulsant, neuroleptic,  
CC vasotropic, cytostatic and gynaecological activities. The polypeptides  
CC and polynucleotides of the invention may be useful for diagnosis,  
CC detection, treatment and/or prevention of disorders of the liver such as  
CC viral, fungal, bacterial or parasitic infections, cirrhosis, Wilson's  
CC disease, gastrointestinal disorders, pancreatic disorders, gallbladder  
CC diseases, immune disorders, blood related disorders, hyperproliferative  
CC disorders, cardiovascular disorders, respiratory disorders,  
CC musculoskeletal system disorders, neurological diseases, endocrine  
CC disorders, reproductive system disorders or developmental and inherited  
CC disorders. The current sequence is that of the human liver-related  
CC protein of the invention. The current sequence is not shown within the  
CC specification per se but was obtained electronically from the USPTO web-  
CC site.

Query Match 58.1%; Score 18; DB 7; Length 140;  
Best Local Similarity 22.2%; Pred. No. 3.4e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 4 CXXXXXXC 12  
Db 104 CTRATATAC 112

RESULT 16  
ADP31453  
ID ADP31453 standard; protein; 144 AA.

AC ADP31453;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2220.

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN WO2004035732-A2.

PD 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 09-JUN-2003; 2003US-047609P.  
PR 09-JUN-2003; 2003US-047641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485225P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-048690P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3451; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present  
XX sequence represents a human secreted protein. The present sequence is  
XX available on WIPOMEB and is not in the specification.

SO Sequence 144 AA;

Query Match 58.1%; Score 18; DB 8; Length 144;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
 XX |  
 DB 7 CATTATTC 15

RESULT 17  
 ID ABO83270 standard; protein; 145 AA.  
 XX ABO83270;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #15445.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US651795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX MPI; 2003-615309/58.

XX N-PSDB; ABD16841.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 32016; 455bp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful in diagnosis and  
 CC therapy of pathological conditions, as molecular targets for diagnostics,  
 CC prophylaxis and treatment of pathological conditions resulting from a  
 CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-P. aeruginosa drugs, as templates for recombinant  
 CC production of P. aeruginosa-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of P. aeruginosa-caused  
 CC infection, and in detection of P. aeruginosa sequences or other sequences  
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html

XX Sequence 145 AA;

Query Match 58.1%; Score 18; DB 7; Length 145;  
 Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
 XX |  
 DB 20 CCASTATTC 28

RESULT 18  
 ID ABO79762 standard; protein; 147 AA.  
 XX ABO79762

AC ABO79762;  
 XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #11937.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US651795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX MPI; 2003-615309/58.

XX N-PSDB; ABD13333.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 28508; 455bp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful in diagnosis and  
 CC therapy of pathological conditions, as molecular targets for diagnostics,  
 CC prophylaxis and treatment of pathological conditions resulting from a  
 CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-P. aeruginosa drugs, as templates for recombinant  
 CC production of P. aeruginosa-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of P. aeruginosa-caused  
 CC infection, and in detection of P. aeruginosa sequences or other sequences  
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html

XX Sequence 147 AA;

Query Match 58.1%; Score 18; DB 7; Length 147;  
 Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
 XX |  
 DB 27 CSTSTTAAC 35

RESULT 19  
 ID ADP31511 standard; protein; 153 AA.  
 XX ADP31511;

XX 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #2278.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 XX cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406589P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476611P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485233P.  
PR 08-JUL-2003; 2003US-0485244P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493373P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Hsieh L, Linnemann T;  
PI Pierce K, Wang Y, Wong JG, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 3509; 428bp; English.  
XX  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases.  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
SQ  
SQ Sequence 153 AA;  
Query Match 58.1%; Score 18; DB 8; Length 153;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXC 12  
Db 79 CTATTTTTC 87  
RESULT 20  
ADN23734  
ID ADN23734 standard; protein; 155 AA.  
XX  
XX ADN23734;  
AC  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polypeptide #6387.  
XX  
KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
XX Bacteria.  
OS  
XX  
XX US2003233675-A1.  
PN  
XX  
XX 18-DEC-2003.  
PD  
XX 20-FEB-2003; 2003US-00369493.  
PF  
XX 21-FEB-2002; 2002US-0360039P.  
XX  
XX (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR NPI; 2004-061375/06.  
XX PT New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
XX PT microbial source, useful for producing plants with improved properties.  
XX  
PS Claim 1; SEQ ID NO 6387; 122pp; English.  
XX  
XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 155 AA;  
Query Match 58.1%; Score 18; DB 8; Length 155;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXC 12  
DB 92 CTTASSSSC 100  
RESULT 21  
ABO83103  
ID ABO83103 standard; protein; 156 AA.  
XX  
XX ABO83103;  
AC  
XX 29-JUL-2004 (first entry)  
DT  
XX  
XX Pseudomonas aeruginosa polypeptide #15278.  
DE  
XX  
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
KW  
XX  
XX Pseudomonas aeruginosa.  
OS  
XX  
XX US651795-B1.  
PN  
XX  
XX 22-APR-2003.  
PD  
XX  
XX 18-FEB-1999; 99US-00252991.  
PF  
XX  
XX 18-FEB-1998; 98US-0074788P.  
PR  
XX 27-JUL-1998; 98US-0094190P.  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
PA  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
PI  
XX WPI; 2003-615309/58.  
DR

DR N-PSDB; ABD16674.  
XX  
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
XX Disclosure; SEQ ID NO 31849; 455pp; English.  
XX  
XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using bioclip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html.  
XX  
SQ Sequence 156 AA;  
Query Match 58.1%; Score 18; DB 7; Length 156;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXC 12  
DB 46 CASAAATAC 54  
RESULT 22  
ADP30759  
ID ADP30759 standard; protein; 156 AA.  
XX  
XX ADP30759;  
AC  
XX 12-AUG-2004 (first entry)  
DT  
XX  
XX Human secreted protein SEQ ID #1526.  
DE  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2004035732-A2.  
PN  
XX  
XX 29-APR-2004.  
PD  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR  
XX 29-AUG-2002; 2002US-0406579P.  
PR  
XX 29-AUG-2002; 2002US-0406585P.  
PR  
XX 29-AUG-2002; 2002US-0406588P.  
PR  
XX 29-AUG-2002; 2002US-0406608P.  
PR  
XX 29-AUG-2002; 2002US-0406611P.  
PR  
XX 29-AUG-2002; 2002US-0406612P.  
PR  
XX 29-AUG-2002; 2002US-0406616P.  
PR  
XX 29-AUG-2002; 2002US-0406640P.  
PR  
XX 29-AUG-2002; 2002US-0406642P.  
PR  
XX 29-AUG-2002; 2002US-0406646P.  
PR  
XX 29-AUG-2002; 2002US-0406653P.  
PR  
XX 29-AUG-2002; 2002US-0406655P.  
PR  
XX 29-AUG-2002; 2002US-0406666P.  
PR  
XX 17-SEP-2002; 2002US-0410947P.  
PR  
XX 17-SEP-2002; 2002US-0410947P.  
PR

PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 09-JUN-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halembeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX MPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2757; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 156 AA;  
Query Match 58.1%; Score 18; DB 8; Length 156;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXC 12  
DB 47 CTATTTTAC 55  
RESULT 23  
ADP31286  
ID ADP31286 standard; protein; 165 AA.  
XX  
AC ADP31286;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2053.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.



06-0001-2003; 200303-04033225F



PR	17-SEP-2002	2002US-0410953P
PR	17-SEP-2002	2002US-0410957P
PR	17-SEP-2002	2002US-0410958P
PR	17-SEP-2002	2002US-0410959P
PR	17-SEP-2002	2002US-0410960P
PR	17-SEP-2002	2002US-0410961P
PR	17-SEP-2002	2002US-0410962P
PR	17-SEP-2002	2002US-0411019P
PR	17-SEP-2002	2002US-0411022P
PR	17-SEP-2002	2002US-0411023P
PR	17-SEP-2002	2002US-0411024P
PR	17-SEP-2002	2002US-0411032P
PR	17-SEP-2002	2002US-0411035P
PR	17-SEP-2002	2002US-0411037P
PR	17-SEP-2002	2002US-0411041P
PR	17-SEP-2002	2002US-0411045P
PR	17-SEP-2002	2002US-0411046P
PR	17-SEP-2002	2002US-0411082P
PR	17-SEP-2002	2002US-0411101P
PR	17-SEP-2002	2002US-0411111P
PR	18-APR-2003	2003US-0463700P
PR	18-APR-2003	2003US-0463708P
PR	18-APR-2003	2003US-0463716P
PR	18-APR-2003	2003US-0463732P
PR	02-MAY-2003	2003US-0467201P
PR	02-MAY-2003	2003US-0467203P
PR	02-MAY-2003	2003US-0467203P
PR	19-MAY-2003	2003US-0471306P
PR	19-MAY-2003	2003US-0471336P
PR	22-MAY-2003	2003US-0472420P
PR	22-MAY-2003	2003US-0472430P
PR	09-JUN-2003	2003US-0476609P
PR	09-JUN-2003	2003US-0476641P
PR	08-JUL-2003	2003US-0485218P
PR	08-JUL-2003	2003US-0485223P
PR	08-JUL-2003	2003US-0485224P
PR	08-JUL-2003	2003US-0485325P
PR	14-JUL-2003	2003US-0486446P
PR	14-JUL-2003	2003US-0486480P
PR	15-JUL-2003	2003US-0486919P
PR	15-JUL-2003	2003US-0486960P
PR	08-AUG-2003	2003US-0493341P
PR	08-AUG-2003	2003US-0493370P
PR	08-AUG-2003	2003US-0493573P
PR	08-AUG-2003	2003US-0493577P
XX		
PA	(FIVE -) FIVE PRIME THERAPEUTICS INC.	
PI	Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;	
PI	Haltenbeck RF, Huang MM, Kothakota S, Halshan L, Linnemann T;	
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;	
XX		
DR	WPI; 2004-348438/32.	
XX		
XX	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
PS	Claim 1; SEQ ID NO 2799; 428bp; English.	
XX		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
CC	available on WIPWEB and is not in the specification.	

XX	Sequence	174 AA;	58.1%;	Score 18;	DB 8;	Length 174;
XX	Query Match		22.2%;	Pred. No. 3.6e+02;		
XX	Best Local Similarity		22.2%;	Pred. No. 3.6e+02;		
XX	Matches	2;	Conservative	0;	Mismatches 7;	Indels 0; Gaps 0
QY	4 CXXXXXXC 12					
DB	122 CTAATTTC 130					
XX	RESULT 27					
XX	ABO73977					
XX	ID ABO73977 standard; protein; 193 AA.					
XX	ABO73977;					
XX	29-JUL-2004 (first entry)					
XX	Pseudomonas aeruginosa polypeptide #6152.					
XX	Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.					
XX	Pseudomonas aeruginosa.					
XX	US6551795-B1.					
XX	22-APR-2003.					
XX	18-FEB-1999; 99US-00252991.					
XX	18-FEB-1998; 98US-0074788P.					
XX	27-JUL-1998; 98US-0094190P.					
XX	(GENO-) GENOME THERAPEUTICS CORP.					
XX	Rubenfield MO, Nolling J, Deloughery C, Bush D;					
XX	WPI; 2003-615309/58.					
XX	N-PSDB; ABD07548.					
XX	Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, of					
XX	useful as molecular targets for diagnostics, prophylaxis and treatment of					
XX	pathological conditions resulting from bacterial infection.					
XX	Disclosure; SEQ ID NO 22723; 455PP; English.					
XX	The invention relates to Pseudomonas aeruginosa polypeptides and the					
XX	polynucleotides encoding them. The sequences are useful in diagnostics and					
XX	therapy of pathological conditions, as molecular targets for diagnostics,					
XX	prophylaxis and treatment of pathological conditions resulting from a					
XX	bacterial infection, for evaluating a compound, such as a polypeptide					
XX	for the ability to bind a P. aeruginosa nucleic acid, as components of					
XX	effective antibacterial targets, as targets for antibacterial drugs,					
XX	including anti-P. aeruginosa drugs, as templates for recombinant					
XX	production of P. aeruginosa-derived peptides or polypeptides, as target					
XX	components for diagnosis and/or treatment of P. aeruginosa-caused					
XX	infection, and in detection of P. aeruginosa sequences or other sequences					
XX	of Pseudomonas species using biotech technology. Sequences ABO67825-					
XX	ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The					
XX	sequence data for this patent did not form part of the printed					
XX	specification but was obtained in electronic format from USPTO at					
XX	seqdata.uspto.gov/sequence.html					
XX	Sequence 193 AA;					
XX	Query Match	58.1%;	Score 18;	DB 7;	Length 193;	
XX	Best Local Similarity	22.2%;	Pred. No. 3.7e+02;			
XX	Matches	2;	Conservative	0;	Mismatches 7;	Indels 0; Gaps 0;
QY	4 CXXXXXXC 12					

Db 61 CTSTSSATC 69

RESULT 28  
ADP30493  
ID ADP30493 standard; protein; 198 AA.  
XX  
AC ADP30493;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1260.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
XX  
PR 29-AUG-2002; 2002US-0406579P.  
XX  
PR 29-AUG-2002; 2002US-0406585P.  
XX  
PR 29-AUG-2002; 2002US-0406588P.  
XX  
PR 29-AUG-2002; 2002US-0406608P.  
XX  
PR 29-AUG-2002; 2002US-0406611P.  
XX  
PR 29-AUG-2002; 2002US-0406612P.  
XX  
PR 29-AUG-2002; 2002US-0406616P.  
XX  
PR 29-AUG-2002; 2002US-0406640P.  
XX  
PR 29-AUG-2002; 2002US-0406642P.  
XX  
PR 29-AUG-2002; 2002US-0406646P.  
XX  
PR 29-AUG-2002; 2002US-0406653P.  
XX  
PR 29-AUG-2002; 2002US-0406656P.  
XX  
PR 17-SEP-2002; 2002US-0410946P.  
XX  
PR 17-SEP-2002; 2002US-0410947P.  
XX  
PR 17-SEP-2002; 2002US-0410948P.  
XX  
PR 17-SEP-2002; 2002US-0410949P.  
XX  
PR 17-SEP-2002; 2002US-0410953P.  
XX  
PR 17-SEP-2002; 2002US-0410957P.  
XX  
PR 17-SEP-2002; 2002US-0410958P.  
XX  
PR 17-SEP-2002; 2002US-0410960P.  
XX  
PR 17-SEP-2002; 2002US-0410961P.  
XX  
PR 17-SEP-2002; 2002US-0410962P.  
XX  
PR 17-SEP-2002; 2002US-0411019P.  
XX  
PR 17-SEP-2002; 2002US-0411022P.  
XX  
PR 17-SEP-2002; 2002US-0411023P.  
XX  
PR 17-SEP-2002; 2002US-0411024P.  
XX  
PR 17-SEP-2002; 2002US-0411032P.  
XX  
PR 17-SEP-2002; 2002US-0411035P.  
XX  
PR 17-SEP-2002; 2002US-0411037P.  
XX  
PR 17-SEP-2002; 2002US-0411041P.  
XX  
PR 17-SEP-2002; 2002US-0411045P.  
XX  
PR 17-SEP-2002; 2002US-0411046P.  
XX  
PR 17-SEP-2002; 2002US-0411048P.  
XX  
PR 17-SEP-2002; 2002US-0411052P.  
XX  
PR 17-SEP-2002; 2002US-0411055P.  
XX  
PR 17-SEP-2002; 2002US-0411073P.  
XX  
PR 17-SEP-2002; 2002US-0411082P.  
XX  
PR 17-SEP-2002; 2002US-0411101P.  
XX  
PR 17-SEP-2002; 2002US-0411111P.  
XX  
PR 18-APR-2003; 2003US-0463700P.  
XX  
PR 18-APR-2003; 2003US-0463708P.  
XX  
PR 18-APR-2003; 2003US-0463716P.  
XX  
PR 18-APR-2003; 2003US-0463732P.  
XX  
PR 02-MAY-2003; 2003US-0467199P.  
XX  
PR 02-MAY-2003; 2003US-0467201P.  
XX  
PR 02-MAY-2003; 2003US-0467203P.  
XX

PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LF, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2491; 428pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX CC encoding a polypeptide which is believed to be cytostatic,  
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX CC composition and methods are useful for diagnosing, preventing and  
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present  
XX CC sequence represents a human secreted protein. The present sequence is  
XX CC available on WIPOMEB and is not in the specification.  
XX  
XX Sequence 198 AA;  
SQ  
Query Match 58.1%; Score 18; DB 8; Length 198;  
Best Local Similarity 22.2%; Pred. No. 3.8e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXC 12  
Db 91 CATTTC 99

RESULT 29  
ADP30492  
ID ADP30492 standard; protein; 198 AA.  
XX  
AC ADP30492;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1259.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX

```
PF 28-AUG-2003; 2003MO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406666P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
PR 17-SEP-2002; 2002US-0410959P.
PR 17-SEP-2002; 2002US-0410960P.
PR 17-SEP-2002; 2002US-0410961P.
PR 17-SEP-2002; 2002US-0410962P.
PR 17-SEP-2002; 2002US-0411019P.
PR 17-SEP-2002; 2002US-0411022P.
PR 17-SEP-2002; 2002US-0411023P.
PR 17-SEP-2002; 2002US-0411024P.
PR 17-SEP-2002; 2002US-0411032P.
PR 17-SEP-2002; 2002US-0411035P.
PR 17-SEP-2002; 2002US-0411037P.
PR 17-SEP-2002; 2002US-0411041P.
PR 17-SEP-2002; 2002US-0411045P.
PR 17-SEP-2002; 2002US-0411046P.
PR 17-SEP-2002; 2002US-0411048P.
PR 17-SEP-2002; 2002US-0411052P.
PR 17-SEP-2002; 2002US-0411055P.
PR 17-SEP-2002; 2002US-0411073P.
PR 17-SEP-2002; 2002US-0411082P.
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-047136P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-048690P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
```

---

```
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linemann T;
PI Pierece K, Wang Y, Wong JGP, Wu G, Khang H;
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
PS Claim 1; SEQ ID NO 2490; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPOMB and is not in the specification.
SQ
XX Sequence 198 AA:
Query Match 58.1%; Score 18; DB 8; Length 198;
Best Local Similarity 22.2%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 4 CXXXXXXC 12
Db 91 CATTYTTTC 99
RESULT 30
ADP0477
ID ADP0477 standard; protein; 198 AA.
XX
XX ADP0477;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1244.
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003MO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
```



PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

## (FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MW, Kochakota S, Haisan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 2479; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and vincinal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.

XX Sequence 198 AA;

Query Match 58.1%; Score 18; DB 8; Length 198;  
Best Local Similarity 22.2%; Pred. No. 3.8e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 12  
DB 54 CATTAATTAC 62

## RESULT 32

ID ADP30491 standard; protein; 198 AA.

XX ADP30491;

XX 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1258.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.

XX OS Homo sapiens.  
XX PN W02004035732-A2.  
XX PD 29-APR-2004.  
XX PF 28-AUG-2003; 2003MO-US026780.

XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 22-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2489; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 198 AA;  
Query Match 58.1%; Score 18; DB 8; Length 198;  
Best Local Similarity 22.2%; Pred. No. 3.8e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
OY 4 CXXXXXXXC 12  
DB 91 CATTTC 99  
RESULT 33  
ADP31281  
ID ADP31281 standard; protein; 228 AA.  
XX  
XX ADP31281;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2048.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406589P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485252P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
XX Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3279; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The



CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.

CC Sequence 228 AA;

Query Match 58.1%; Score 18; DB 8; Length 228;

Best Local Similarity 22.2%; Pred. No. 3.9e+02; Mismatches 7; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

4 CXXXXXXXXC 12

Db 183 CTTTAATC 191

RESULT 34

ABO80501

ID ABO80501 standard; protein; 229 AA.

XX ABO80501;

XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #12676.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US651795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR MPI; 2003-615309/58.

DR N-PSDB; ABD14072.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 29247; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using biotech technology. Sequences ABO67826-  
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html

XX Sequence 229 AA;

Query Match 58.1%; Score 18; DB 7; Length 229;

Best Local Similarity 22.2%; Pred. No. 3.9e+02; Mismatches 7; Indels 0; Gaps 0;

4 CXXXXXXXXC 12

Db 155 CATTATAC 163

RESULT 35

ABO77555

ID ABO77555 standard; protein; 233 AA.

XX ABO77555;

XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #9730.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US651795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR MPI; 2003-615309/58.

DR N-PSDB; ABD11126.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 26301; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using biotech technology. Sequences ABO67826-  
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html

XX Sequence 233 AA;

Query Match 58.1%; Score 18; DB 7; Length 233;

Best Local Similarity 22.2%; Pred. No. 3.9e+02; Mismatches 7; Indels 0; Gaps 0;

4 CXXXXXXXXC 12

Db 43 CSTASASC 51

RESULT 36

ADP31468  
ID ADP31468 standard; protein; 234 AA.  
XX  
AC ADP31468;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2235.  
XX  
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406589P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411010P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486466P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;  
PI Halenbeck RP, Huang MM, Kochakota S, Hsieh L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3466; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic, the  
CC antinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMER and is not in the specification.  
XX  
XX Sequence 234 AA;  
SQ  
Query Match 58.1%; Score 18; DB 8; Length 234;  
Best Local Similarity 22.2%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXX 12  
Db 213 CTTATTAC 221  
RESULT 37  
ADP31616  
ID ADP31616 standard; protein; 242 AA.  
XX  
AC ADP31616;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2383.  
XX  
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF 29-AUG-2002; 2002US-0406576P.  
XX  
PR 29-AUG-2002; 2002US-0406579P.



XX	03-JAN-2002.
PD	
XX	
PF	20-JUN-2001; 2001WO-US019665.
XX	
XX	28-JUN-2000; 2000US-0214392P.
PR	19-SEP-2000; 2000US-0233519P.
PR	18-APR-2001; 2001US-0264438P.
PR	18-APR-2001; 2001US-0264617P.
PR	18-APR-2001; 2001US-0284730P.
XX	
PA	(BRIM ) BRISTOL-MYERS SQUIBB CO.
XX	
PI	Saiwati ME, Gottardis MM, Krystek SR, Attar RM, Sack JS;
DR	WPI; 2002-164433/21.
XX	
XX	New selective androgen receptor modulator for treating tumors, comprises
PT	antagonist activity in a hormone-dependent tumor and no or agonist
PT	activity against other, non-tumor tissues containing the receptor.
XX	
PS	Disclosure; Page 19; 140pp; English.
XX	
CC	The invention describes a selective androgen receptor modulator (SARM)
CC	(I) which exhibits antagonist activity in a hormone-dependent tumour
CC	while exhibiting no activity or agonist activity against other non-tumour
CC	tissues containing the androgen receptor. (I) is useful for inhibiting
CC	the growth of hormone-dependent tumour cells (prostate tumour cells) in a
CC	patient. (II) is also useful for treating a condition selected from
CC	hyperandrogenism, acne, seborrhea, Alzheimer's disease, androgenic alopecia,
CC	hypogonadism, hypopituitarism, benign prostatic hyperplasia, adenomas or
CC	neoplasias of the prostate, treatment of benign or malignant tumour cells
CC	containing the androgen receptor, pancreatic cancers, modulation of
CC	vascular endothelial growth factor (VEGF) expression for use as
CC	antiangiogenic agents, osteoporosis, suppressing spermatogenesis, libido,
CC	cachexia, endometriosis, polycystic ovary syndrome, anorexia, androgen
CC	dependent age-related diseases and conditions, male menopause, male
CC	hormone replacement, male and female sexual dysfunction, and inhibition
CC	of muscular atrophy in ambulatory patients. (I) is also useful for
CC	treating muscle wasting and gynaecomastia. This sequence represents the
CC	androgen receptor ligand binding domain (AR-LBD), described in the method
CC	of the invention
XX	
XX	Sequence 246 AA;
QQ	
Query Match	58.1%; Score 18; DB 5; Length 246;
Best Local Similarity	22.2%; Pred. No. 4e+02;
Matches	2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY	4 CXXXXXXXC 12
Db	173 CAAAAAASC 181
ABO74848	
ID	ABO74848 standard; protein; 246 AA.
AC	ABO74848;
XX	
DT	29-JUL-2004 (first entry)
XX	
DE	Pseudomonas aeruginosa polypeptide #7023.
XX	
KW	Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX	
OS	Pseudomonas aeruginosa.
XX	
PN	US6551795-B1.
XX	
PD	22-APR-2003.
XX	
PF	18-FEB-1999; 99US-00252991.

XX	18-FEB-1998;	98US-0074788P.	
PR	27-JUL-1998;	98US-0094190P.	
PA	(GENO-)	GENOME THERAPEUTICS CORP.	
PI	Rubenfield MJ, Nolling J, Deloughery C, Bush D,		
DR	WPI; 2003-615309/58.		
XX	N-PSDB; A8D06419.		
PT	Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,		
PP	useful as molecular targets for diagnostics, prophylaxis and treatment of		
XX	pathological conditions resulting from bacterial infection.		
PS	Disclosure; SEQ ID NO 23594; 455pp; English.		
XX			
CC	The invention relates to Pseudomonas aeruginosa polypeptides and the		
CC	polynucleotides encoding them. The sequences are useful in diagnosis and		
CC	therapy of pathological conditions, as molecular targets for diagnostics,		
CC	prophylaxis and treatment of pathological conditions resulting from a		
CC	bacterial infection, for evaluating a compound, such as a polypeptide,		
CC	for the ability to bind a P. aeruginosa nucleic acid, as components of		
CC	effective antibacterial targets, as targets for antibacterial drugs,		
CC	including anti-P. aeruginosa drugs, as templates for recombinant		
CC	production of P. aeruginosa-derived peptides or polypeptides, as target		
CC	components for diagnosis and/or treatment of P. aeruginosa-caused		
CC	infection, and in detection of P. aeruginosa sequences or other sequences		
CC	of Pseudomonas species using biochip technology. Sequences A8067826-		
CC	A8084336 represent P. aeruginosa polypeptides of the invention. Note: The		
CC	sequence data for this patent did not form part of the printed		
CC	specification but was obtained in electronic format from USPTO at		
CC	seqdata.uspto.gov/sequence.html		
XX			
SQ	Sequence 246 AA;		
XX			
Query Match	58.1%;	Score 18;	DB 7;
Best Local Similarity	22.2%;	Pred. NO. 4e+02;	Length 246;
Matches	2;	Conservative	0;
		Mismatches	7;
		Indels	0;
		Gaps	0
QY	4 CXXXXXXXC 12		
DB	202 CSTSSATC 210		
RESULT 40			
ADP30754			
ID	ADP30754	standard; protein; 249 AA.	
XX			
AC	ADP30754;		
XX			
DT	12-AUG-2004	(first entry)	
XX			
DE	Human secreted protein SEQ ID #1521.		
XX			
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;		
XX	cancer; inflammatory; immune; human secreted protein.		
OS	Homo sapiens.		
XX			
PN	WO2004035732-A2.		
XX			
PD	29-APR-2004.		
XX			
PF	28-AUG-2003;	2003WO-US026780.	
XX			
XX	29-AUG-2002;	2002US-0406576P.	
PR	29-AUG-2002;	2002US-0406579P.	
PR	29-AUG-2002;	2002US-0406585P.	
PR	29-AUG-2002;	2002US-0406588P.	
PR	29-AUG-2002;	2002US-0406608P.	
PR	29-AUG-2002;	2002US-0406611P.	
PR	29-AUG-2002;	2002US-0406612P.	

PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485255P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halsebeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX MPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

XX  
PS Claim 1; SEQ ID NO 2752; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antitumorigenic, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 249 AA;  
XX  
Query Match 58.1%; Score 18; DB 8; Length 249;  
Best Local Similarity 22.2%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXC 12  
DB 238 CTTAAAAAC 246  
XX  
RESULT 41  
ADP31396  
ID ADP31396 standard; protein; 254 AA.  
XX  
AC ADP31396;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DB Human secreted protein SEQ ID #2163.  
XX  
KW Cytostatic; Antitumorigenic; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN MO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.

XX	17-SEP-2002;	2002US-0411032P.
PR	17-SEP-2002;	2002US-0411035P.
PR	17-SEP-2002;	2002US-0411037P.
PR	17-SEP-2002;	2002US-0411041P.
PR	17-SEP-2002;	2002US-0411045P.
PR	17-SEP-2002;	2002US-0411046P.
PR	17-SEP-2002;	2002US-0411048P.
PR	17-SEP-2002;	2002US-0411052P.
PR	17-SEP-2002;	2002US-0411055P.
PR	17-SEP-2002;	2002US-0411073P.
PR	17-SEP-2002;	2002US-0411082P.
PR	17-SEP-2002;	2002US-0411101P.
PR	17-SEP-2002;	2002US-0411111P.
PR	18-APR-2003;	2003US-0463700P.
PR	18-APR-2003;	2003US-0463708P.
PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US-0463732P.
PR	02-MAY-2003;	2003US-0467199P.
PR	02-MAY-2003;	2003US-0467201P.
PR	02-MAY-2003;	2003US-0467203P.
PR	02-MAY-2003;	2003US-0467230P.
PR	19-MAY-2003;	2003US-0471306P.
PR	19-MAY-2003;	2003US-0471336P.
PR	22-MAY-2003;	2003US-0472420P.
PR	22-MAY-2003;	2003US-0472430P.
PR	09-JUN-2003;	2003US-0476609P.
PR	09-JUN-2003;	2003US-0476641P.
PR	08-JUL-2003;	2003US-0485218P.
PR	08-JUL-2003;	2003US-0485223P.
PR	08-JUL-2003;	2003US-0485224P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486446P.
PR	14-JUL-2003;	2003US-0486480P.
PR	15-JUL-2003;	2003US-0486960P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.
XX	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
PA	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;	
PI	Hallenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;	
PI	Pierce K, Wang Y, Wong JCF, Wu G, Zhang H;	
DR	WPI; 2004-348438/32.	
XX	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX	Claim 1; SEQ ID NO 3394; 428bp; English.	
XX	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
CC	available on WIPOMB and is not in the specification.	
XX	Sequence 254 AA;	
XX		

Query Match	Similarity	Score 18;	DB 8;	Length 254;
Best Local	Similarity 22.28;	Pred. NC 4e+02;		
Matches 2;	Conservative	0;	Mismatches 7;	Indels 0;
Gaps 0;				
QY	4 CXXXXXXC	12		
DB	77 CTATTAAC	85		

RESULT 42	
ADP30846	
ID	ADP30846 standard; protein; 279 AA.
XX	
AC	ADP30846;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Human secreted protein SEQ ID #1613.
XX	
KW	Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide
KW	cancer; Inflammatory; Immune; human secreted protein.
XX	
OS	Homo sapiens.
XX	
PN	WO2004035732-A2.
XX	
PD	29-APR-2004.
XX	
PF	28-AUG-2003; 2003WO-US026780.
XX	
PR	29-AUG-2002; 2002US-0406576P.
PR	29-AUG-2002; 2002US-0406579P.
PR	29-AUG-2002; 2002US-0406585P.
PR	29-AUG-2002; 2002US-0406588P.
PR	29-AUG-2002; 2002US-0406608P.
PR	29-AUG-2002; 2002US-0406611P.
PR	29-AUG-2002; 2002US-0406612P.
PR	29-AUG-2002; 2002US-0406616P.
PR	29-AUG-2002; 2002US-0406640P.
PR	29-AUG-2002; 2002US-0406642P.
PR	29-AUG-2002; 2002US-0406646P.
PR	29-AUG-2002; 2002US-0406653P.
PR	29-AUG-2002; 2002US-0406655P.
PR	29-AUG-2002; 2002US-0406666P.
PR	17-SEP-2002; 2002US-0410944P.
PR	17-SEP-2002; 2002US-0410947P.
PR	17-SEP-2002; 2002US-0410948P.
PR	17-SEP-2002; 2002US-0410949P.
PR	17-SEP-2002; 2002US-0410953P.
PR	17-SEP-2002; 2002US-0410957P.
PR	17-SEP-2002; 2002US-0410958P.
PR	17-SEP-2002; 2002US-0410959P.
PR	17-SEP-2002; 2002US-0410960P.
PR	17-SEP-2002; 2002US-0410961P.
PR	17-SEP-2002; 2002US-0410962P.
PR	17-SEP-2002; 2002US-0411019P.
PR	17-SEP-2002; 2002US-0411022P.
PR	17-SEP-2002; 2002US-0411023P.
PR	17-SEP-2002; 2002US-0411024P.
PR	17-SEP-2002; 2002US-0411032P.
PR	17-SEP-2002; 2002US-0411035P.
PR	17-SEP-2002; 2002US-0411037P.
PR	17-SEP-2002; 2002US-0411041P.
PR	17-SEP-2002; 2002US-0411045P.
PR	17-SEP-2002; 2002US-0411046P.
PR	17-SEP-2002; 2002US-0411048P.
PR	17-SEP-2002; 2002US-0411052P.
PR	17-SEP-2002; 2002US-0411055P.
PR	17-SEP-2002; 2002US-0411073P.
PR	17-SEP-2002; 2002US-0411082P.
PR	17-SEP-2002; 2002US-0411101P.
PR	17-SEP-2002; 2002US-0411111P.
PR	18-APR-2003; 2003US-0463700P.
PR	18-APR-2003; 2003US-0463708P.
PR	18-APR-2003; 2003US-0463716P.
PR	18-APR-2003; 2003US-0463732P.
PR	02-MAY-2003; 2003US-0467199P.
PR	02-MAY-2003; 2003US-0467201P.
PR	02-MAY-2003; 2003US-0467203P.
PR	02-MAY-2003; 2003US-0467205P.
PR	19-MAY-2003; 2003US-0471306P.



PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX	WPI; 2004-348438/32.
XX	
XX	New nucleic acid molecule for diagnosing, preventing or treating diseases
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT	genetic, bacterial and viral diseases.
XX	
PS	Claim 1; SEQ ID NO 3521, 428pp; English.
XX	
CC	The present invention relates to an isolated nucleic acid molecule
CC	encoding a polypeptide which is believed to be cytostatic,
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC	composition and methods are useful for diagnosing, preventing and
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,
CC	immune, metabolic, genetic, bacterial and viral diseases. The present
CC	sequence represents a human secreted protein. The present sequence is
CC	available on WIPOMEB and is not in the specification.
XX	
SO	Sequence 279 AA;
Oy	4 CXXXXXXXXX 12
Db	66 CATTAATC 74
RESULT 44	
ID	ADP30844
XX	ADP30844 standard; protein; 279 AA.
AC	ADP30844;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Human secreted protein SEQ ID #1611.
XX	
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW	cancer; inflammatory; immune; human secreted protein.
XX	
OS	Homo sapiens.
XX	
PN	WO2004035732-A2.
XX	
PD	29-APR-2004.
XX	
PF	28-AUG-2003; 2003WO-US026780.
XX	
PR	29-AUG-2002; 2002US-0406576P.
PR	29-AUG-2002; 2002US-0406579P.
PR	29-AUG-2002; 2002US-0406585P.
PR	29-AUG-2002; 2002US-0406588P.
PR	29-AUG-2002; 2002US-0406608P.
PR	29-AUG-2002; 2002US-0406611P.
PR	29-AUG-2002; 2002US-0406612P.
PR	29-AUG-2002; 2002US-0406616P.
PR	29-AUG-2002; 2002US-0406642P.
PR	29-AUG-2002; 2002US-0406644P.
PR	29-AUG-2002; 2002US-0406646P.
PR	29-AUG-2002; 2002US-0406653P.
PR	29-AUG-2002; 2002US-0406655P.
PR	29-AUG-2002; 2002US-0406666P.
PR	17-SEP-2002; 2002US-0410946P.
PR	17-SEP-2002; 2002US-0410947P.
PR	17-SEP-2002; 2002US-0410948P.
PR	17-SEP-2002; 2002US-0410949P.
PR	17-SEP-2002; 2002US-0410953P.
PR	17-SEP-2002; 2002US-0410957P.
PR	17-SEP-2002; 2002US-0410958P.
PR	17-SEP-2002; 2002US-0410959P.

XX	17-SEP-2002; 2002US-0410961P.	PR
XX	17-SEP-2002; 2002US-0410961P.	PR
XX	17-SEP-2002; 2002US-0410962P.	PR
XX	17-SEP-2002; 2002US-0410962P.	PR
XX	17-SEP-2002; 2002US-0411019P.	PR
XX	17-SEP-2002; 2002US-0411022P.	PR
XX	17-SEP-2002; 2002US-0411023P.	PR
XX	17-SEP-2002; 2002US-0411024P.	PR
XX	17-SEP-2002; 2002US-0411032P.	PR
XX	17-SEP-2002; 2002US-0411035P.	PR
XX	17-SEP-2002; 2002US-0411037P.	PR
XX	17-SEP-2002; 2002US-0411041P.	PR
XX	17-SEP-2002; 2002US-0411045P.	PR
XX	17-SEP-2002; 2002US-0411046P.	PR
XX	17-SEP-2002; 2002US-0411048P.	PR
XX	17-SEP-2002; 2002US-0411052P.	PR
XX	17-SEP-2002; 2002US-0411055P.	PR
XX	17-SEP-2002; 2002US-0411073P.	PR
XX	17-SEP-2002; 2002US-0411082P.	PR
XX	17-SEP-2002; 2002US-0411101P.	PR
XX	17-SEP-2002; 2002US-0411111P.	PR
XX	18-APR-2003; 2003US-0463700P.	PR
XX	18-APR-2003; 2003US-0463708P.	PR
XX	18-APR-2003; 2003US-0463716P.	PR
XX	18-APR-2003; 2003US-0463732P.	PR
XX	02-MAY-2003; 2003US-0467199P.	PR
XX	02-MAY-2003; 2003US-0467201P.	PR
XX	02-MAY-2003; 2003US-0467203P.	PR
XX	02-MAY-2003; 2003US-0467230P.	PR
XX	19-MAY-2003; 2003US-0471306P.	PR
XX	19-MAY-2003; 2003US-0471336P.	PR
XX	22-MAY-2003; 2003US-0472420P.	PR
XX	22-MAY-2003; 2003US-0472430P.	PR
XX	09-JUN-2003; 2003US-0476609P.	PR
XX	09-JUN-2003; 2003US-0476641P.	PR
XX	08-JUL-2003; 2003US-0485218P.	PR
XX	08-JUL-2003; 2003US-0485223P.	PR
XX	08-JUL-2003; 2003US-0485224P.	PR
XX	08-JUL-2003; 2003US-0485325P.	PR
XX	14-JUL-2003; 2003US-0486446P.	PR
XX	14-JUL-2003; 2003US-0486480P.	PR
XX	15-JUL-2003; 2003US-0486891P.	PR
XX	15-JUL-2003; 2003US-0486896P.	PR
XX	08-AUG-2003; 2003US-0493341P.	PR
XX	08-AUG-2003; 2003US-0493370P.	PR
XX	08-AUG-2003; 2003US-0493573P.	PR
XX	08-AUG-2003; 2003US-0493577P.	PR
XX	(FIVE-) FIVE PRIME THERAPEUTICS INC.	PA
XX	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;	PI
XX	Halebeck RT, Huang MM, Kothakota S, Haislan L, Linnemann T;	PI
XX	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;	PI
XX	WPI; 2004-348438/32.	DR
XX	New nucleic acid molecule for diagnosing, preventing or treating diseases	PT
XX	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	PT
XX	genetic, bacterial and viral diseases.	PT
XX	Claim 1; SEQ ID NO 2842; 428bp; English.	PS
XX	The present invention relates to an isolated nucleic acid molecule	CC
XX	encoding a polypeptide which is believed to be cytostatic,	CC
XX	antitumorigenic, immunosuppressive, antibacterial and virucidal. The	CC
XX	composition and methods are useful for diagnosing, preventing and	CC
XX	treating diseases such as proliferative (e.g. cancer), inflammatory,	CC
XX	immune, metabolic, genetic, bacterial and viral diseases. The present	CC
XX	sequence represents a human secreted protein. The present sequence is	CC
XX	available on WIPOMB and is not in the specification.	CC
XX	Sequence 279 AA;	XX



Best Local Similarity 22.2%; Pred. No. 4.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 247 CTATTTAAC 255

## RESULT 45

AB48050 standard; protein; 281 AA.

AC AAB48050;  
XX  
DT 19-MAR-2001 (first entry)  
XX  
DE Signal transduction protein.  
XX  
XX Zea mays; maize; signal transduction protein; phytohormone; ethylene;  
KM auxin; cytokinin; gibberellin; immunogen.  
XX

OS Zea mays.

PN WO200070059-A2.

PD 23-NOV-2000.

PF 28-APR-2000; 2000MO-US011687.

PR 14-MAY-1999; 99US-0134292P.

PR 08-JUL-1999; 99US-0142996P.

PA (PION-) PIONEER HI-BRED INT. INC.

PI Helentjaris TG;

DR WPI; 2001-031929/04.

DR N-PSDB; AAC64271.

PT New signal transduction nucleic acids and encoded proteins useful for  
PT regulating phytohormone expression, including ethylene, auxin,  
PT cytokinins and gibberellin, to provide control of plant response to  
PT environmental stresses.

PS Claim 13; Page 102; 126pp; English.

XX The invention provides Zea mays signal transduction proteins and encoding  
XX nucleotide sequences. The nucleic acids are useful for regulating  
XX expression of phytohormones, including ethylene, auxin, cytokinins, and  
XX gibberellin, to effect developmental changes in plants and provide  
XX control of plant response to environmental stresses. They may also be  
XX used as probes or amplification primers in the detection, quantitation or  
XX isolation of gene transcripts, for detecting mutations in the gene, for  
XX monitoring upregulation of expression or changes in enzyme activity in  
XX screening assays of compounds, for detection of any number of allelic  
XX variants, or for site-directed mutagenesis in eukaryotic cells. They may  
XX further be used for recombinant expression of their encoded polypeptides,  
XX as immunogens in the preparation or screening of antibodies, and in sense  
XX or antisense suppression of genes in a host cell, tissue or plant. The  
XX proteins may be used in assays for enzyme agonists or antagonists, as  
XX immunogens or antigens to obtain antibodies specifically immunoreactive  
XX with the proteins. The present sequence represents a signal transduction  
XX protein of the invention

SQ Sequence 281 AA;

Query Match 58.1%; Score 18; DB 4; Length 281;

Best Local Similarity 22.2%; Pred. No. 4.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 11 CSATTTSSC 19

RESULT 46...  
ADM43729 standard; protein; 281 AA.

AC ADM43729;

DT 24-MAR-2005 (first entry)

DE Corn DREB-type transcription factor, seq id 2.

XX Drought resistance; crop improvement; bold tolerance;  
KM transcription factor; DREB.

OS Zea mays.

PN CN1472222-A.

PD 04-FEB-2004.

PF 29-JUL-2002; 2002CN-00125372.

PR 29-JUL-2002; 2002CN-00125372.

PA (UYQI ) UNIV QINGHUA.

PI Liu Q, Qin F, Zhao J;

DR WPI; 2004-317379/30.

DR N-PSDB; ADM43728.

PT DREB transcription factor of corn and its encoding genes and use.

PS Disclosure; SEQ ID NO 2; 18pp; Chinese.

XX The present invention discloses 2 corn DREB-type transcription factors,  
XX CC MADREB1 and MADREB2, which are important in conditions of cold and  
XX drought. Also disclosed are their coding genes and their application in  
XX culture the cold and drought resistant plants. The current sequence  
XX represents a DREB-type transcription factor.

SQ Sequence 281 AA;

Query Match 58.1%; Score 18; DB 8; Length 281;

Best Local Similarity 22.2%; Pred. No. 4.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 11 CSATTTSSC 19

## RESULT 47

AB079000 standard; protein; 286 AA.

AC AB079000;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #1175.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

XX



PS Claim 1; SEQ ID NO 3190; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule

CC encoding a polypeptide which is believed to be cytostatic,

CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is

CC available on WIPOMER and is not in the specification.

XX

SQ Sequence 297 AA;

Query Match 58.1%; Score 18; DB 8; Length 297;

Best Local Similarity 22.2%; Pred. No. 4.2e+02; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7;

QY 4 CXXXXXXC 12

DB 39 CAATTATAC 47

RESULT 49

ID ADP31454 standard; protein; 312 AA.

XX ADP31454;

AC 12-AUG-2004 (first entry)

DT 12-AUG-2004 (first entry)

XX

DE Human secreted protein SEQ ID #2221.

XX

KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

XX cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

XX

PN MO2004035732-AA2.

XX

PD 29-APR-2004.

XX

PF 28-AUG-2003; 2003WO-US026780.

XX

PR 29-AUG-2002; 2002US-0406576P.

XX

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406583P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406640P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410961P.

PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.

PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.

PR 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.

PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.

PR 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411011P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485252P.

PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;

PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

XX

DR WPI; 2004-348438/32.

XX

PT New nucleic acid molecule for diagnosing, preventing or treating diseases

PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,

PT genetic, bacterial and viral diseases.

XX

PS Claim 1; SEQ ID NO 3452; 428bp; English.

XX

XX The present invention relates to an isolated nucleic acid molecule

CC encoding a polypeptide which is believed to be cytostatic,

CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is

CC available on WIPOMER and is not in the specification.

XX

SQ Sequence 312 AA;

Query Match 58.1%; Score 18; DB 8; Length 312;

Best Local Similarity 22.2%; Pred. No. 4.2e+02; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7;

QY 4 CXXXXXXC 12

DB 7 CAATTATTC 15

```
RESULT 50
ADP31615
ID ADP31615 standard; protein; 317 AA.
XX
AC ADP31615;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2382.
XX
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
XX
PR 29-AUG-2002; 2002US-0406585P.
XX
PR 29-AUG-2002; 2002US-0406588P.
XX
PR 29-AUG-2002; 2002US-0406589P.
XX
PR 29-AUG-2002; 2002US-0406608P.
XX
PR 29-AUG-2002; 2002US-0406611P.
XX
PR 29-AUG-2002; 2002US-0406612P.
XX
PR 29-AUG-2002; 2002US-0406616P.
XX
PR 29-AUG-2002; 2002US-0406640P.
XX
PR 29-AUG-2002; 2002US-0406642P.
XX
PR 29-AUG-2002; 2002US-0406646P.
XX
PR 29-AUG-2002; 2002US-0406653P.
XX
PR 29-AUG-2002; 2002US-0406655P.
XX
PR 29-AUG-2002; 2002US-0406666P.
XX
PR 17-SEP-2002; 2002US-0410946P.
XX
PR 17-SEP-2002; 2002US-0410947P.
XX
PR 17-SEP-2002; 2002US-0410948P.
XX
PR 17-SEP-2002; 2002US-0410949P.
XX
PR 17-SEP-2002; 2002US-0410953P.
XX
PR 17-SEP-2002; 2002US-0410957P.
XX
PR 17-SEP-2002; 2002US-0410958P.
XX
PR 17-SEP-2002; 2002US-0410959P.
XX
PR 17-SEP-2002; 2002US-0410960P.
XX
PR 17-SEP-2002; 2002US-0410961P.
XX
PR 17-SEP-2002; 2002US-0410962P.
XX
PR 17-SEP-2002; 2002US-0410966P.
XX
PR 17-SEP-2002; 2002US-0411019P.
XX
PR 17-SEP-2002; 2002US-0411022P.
XX
PR 17-SEP-2002; 2002US-0411023P.
XX
PR 17-SEP-2002; 2002US-0411024P.
XX
PR 17-SEP-2002; 2002US-0411024P.
XX
PR 17-SEP-2002; 2002US-0411032P.
XX
PR 17-SEP-2002; 2002US-0411035P.
XX
PR 17-SEP-2002; 2002US-0411037P.
XX
PR 17-SEP-2002; 2002US-0411041P.
XX
PR 17-SEP-2002; 2002US-0411045P.
XX
PR 17-SEP-2002; 2002US-0411046P.
XX
PR 17-SEP-2002; 2002US-0411048P.
XX
PR 17-SEP-2002; 2002US-0411052P.
XX
PR 17-SEP-2002; 2002US-0411055P.
XX
PR 17-SEP-2002; 2002US-0411073P.
XX
PR 17-SEP-2002; 2002US-0411082P.
XX
PR 17-SEP-2002; 2002US-0411101P.
XX
PR 17-SEP-2002; 2002US-0411111P.
XX
PR 18-APR-2003; 2003US-0463700P.
XX
PR 18-APR-2003; 2003US-0463708P.
XX
PR 18-APR-2003; 2003US-0463716P.
XX
PR 18-APR-2003; 2003US-0463732P.
XX
PR 02-MAY-2003; 2003US-0467199P.
XX
PR 02-MAY-2003; 2003US-0467201P.
XX
PR 02-MAY-2003; 2003US-0467203P.
XX
PR 19-MAY-2003; 2003US-0471306P.
XX
PR 19-MAY-2003; 2003US-0471336P.
```

```
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485225P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR MPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3613; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
SQ Sequence 317 AA;
XX
Query Match 58.1%; Score 18; DB 8; Length 317;
Best Local Similarity 22.2%; Pred. No. 4.2e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 4 CXXXXXXXC 12
Db 145 CATTATATC 153
```

Search completed: January 4, 2006, 15:55:50  
Job time : 115.652 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 15.5217 Seconds  
(without alignments)  
92.983 Million cell updates/sec

Title: US-09-932-322-3

Perfect score: 31  
Sequence: 1 XXXXXXXXXXXXXXXX 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 1000 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	58.1	155	2 T28945	hypothetical prote
2	18	58.1	470	2 T45962	hypothetical prote
3	18	58.1	1237	2 A34598	ecdysone-induced p
4	18	58.1	1394	2 B34598	ecdysone-induced p
5	18	58.1	1443	2 S05979	steroid hormone re
6	17	54.8	199	2 T33355	hypothetical prote
7	17	54.8	212	2 S63342	hypothetical prote
8	17	54.8	214	2 C70812	probable lpgq prot
9	17	54.8	244	2 H72532	hypothetical prote
10	17	54.8	244	2 T40415	hypothetical prote
11	17	54.8	259	2 H71353	hypothetical prote
12	17	54.8	281	2 AC0661	conserved hypothet
13	17	54.8	281	2 AC0661	probable lipoprote
14	17	54.8	293	2 T48975	xyloglucan endo-tr
15	17	54.8	298	2 T01948	hypothetical prote
16	17	54.8	315	2 A85019	probable bHLH DNA-
17	17	54.8	373	2 AB1224	cobalamin biosynth
18	17	54.8	373	2 AD1577	cobalamin biosynth
19	17	54.8	415	2 T43352	nuclear receptor N
20	17	54.8	507	2 T33024	hypothetical prote
21	17	54.8	530	2 T48004	multifunctional am
22	17	54.8	533	2 A42143	skn-1 - Caenorhabd
23	17	54.8	575	2 T07739	probable ferriochel
24	17	54.8	582	2 T07953	chromomodulin pre
25	17	54.8	582	2 T07953	lectin-like protei
26	17	54.8	584	2 B25682	lectin-like protei
27	17	54.8	584	2 B25682	homeotic protein B
28	17	54.8	589	2 S41011	homeotic protein B
29	17	54.8	595	2 A42086	CD30 antigen precu
30	17	54.8	600	2 D83286	hypothetical prote

30	17	54.8	617	2 T15408	hypothetical prote
31	17	54.8	624	2 A36962	laccase (EC 1.10.3
32	17	54.8	769	2 A41029	integrin beta-8 ch
33	17	54.8	791	2 A46140	diacylglycerol kin
34	17	54.8	796	2 B46140	diacylglycerol kin
35	17	54.8	799	2 A38308	integrin beta-5 ch
36	17	54.8	934	2 T05201	hypothetical prote
37	17	54.8	938	2 T41932	hypothetical prote
38	17	54.8	195	2 T15881	hypothetical prote
39	17	54.8	2395	2 S50820	surface protein ty
40	17	54.8	2704	2 S09118	G surface protein
41	17	54.8	2718	2 A23475	G surface protein
42	17	54.8	8243	2 T31307	type I fatty acid
43	16	51.6	52	2 G82010	probable lipoprote
44	16	51.6	87	2 H66753	prophage pi2 prote
45	16	51.6	91	2 US0036	Clara cell 10K pro
46	16	51.6	96	1 UGMS	uteroglobin precu
47	16	51.6	96	2 A36581	polychlorinated bi
48	16	51.6	98	2 A10334	conserved hypothet
49	16	51.6	103	2 S76421	ferredoxin [pre-25
50	16	51.6	104	2 J00863	hypothetical 11.6K
51	16	51.6	116	2 S09845	hypothetical prote
52	16	51.6	121	2 A44544	amine dehydrogenas
53	16	51.6	128	2 S74085	lutropin beta chal
54	16	51.6	138	2 T21792	hypothetical prote
55	16	51.6	148	2 B72513	hypothetical prote
56	16	51.6	152	2 H97405	hypothetical prote
57	16	51.6	153	2 P75478	conserved hypothet
58	16	51.6	156	2 T31839	hypothetical prote
59	16	51.6	158	2 A61091	lutropin beta chal
60	16	51.6	159	2 I51373	lutinizing hormon
61	16	51.6	160	1 S17194	endothelin 2 precu
62	16	51.6	166	2 I51242	lutinizing hormon
63	16	51.6	166	2 S69590	hypothetical prote
64	16	51.6	182	2 T01909	transcripton init
65	16	51.6	190	2 S48101	xyloglucan endo-1,
66	16	51.6	195	2 T19151	hypothetical prote
67	16	51.6	207	2 A69446	conserved hypothet
68	16	51.6	209	2 S51480	drought-induced pr
69	16	51.6	238	2 T04166	thumatin-like pro
70	16	51.6	246	2 AG3519	enterobactin synth
71	16	51.6	246	2 T20341	hypothetical prote
72	16	51.6	251	2 T21786	hypothetical prote
73	16	51.6	251	2 A55523	hypothetical prote
74	16	51.6	252	2 T04739	hypothetical prote
75	16	51.6	255	2 T38426	lymphocyte activat
76	16	51.6	260	2 T23033	hypothetical prote
77	16	51.6	261	2 C64049	hypothetical prote
78	16	51.6	261	2 S63604	homobox protein G
79	16	51.6	262	2 S01914	B1 membrane glycop
80	16	51.6	265	2 H97280	hydroxyethylthiaz
81	16	51.6	266	2 T02807	arsenate reductase
82	16	51.6	266	2 H84409	thiamin biosynthal
83	16	51.6	268	2 S19184	hypothetical prote
84	16	51.6	268	2 H81806	histidine-binding
85	16	51.6	268	2 B81063	amino acid ABC tra
86	16	51.6	269	2 T35647	phosphomethylpyrim
87	16	51.6	275	2 T32813	hypothetical prote
88	16	51.6	276	2 T01177	hypothetical prote
89	16	51.6	276	2 T52349	F-box protein FBI2
90	16	51.6	277	2 A46241	interferon respons
91	16	51.6	278	2 T39749	hypothetical prote
92	16	51.6	279	2 T51017	hypothetical prote
93	16	51.6	286	2 G97442	hypothetical prote
94	16	51.6	289	2 H75551	xyloglucan endo-1,
95	16	51.6	293	2 T31840	probable transcript
96	16	51.6	295	2 S48102	hypothetical prote
97	16	51.6	299	2 F84785	apoptosis-mediatin
98	16	51.6	303	2 T23321	conserved hypothet
99	16	51.6	309	2 T23321	hypothetical prote
100	16	51.6	327	2 A46484	hypothetical prote
101	16	51.6	331	2 A71338	hypothetical prote
102	16	51.6	332	2 T26339	hypothetical prote

103	16	51.6	333	2	A96829	probable RING fing
104	16	51.6	336	2	S76979	hypothetical prote
105	16	51.6	341	2	S63666	platelet activatin
106	16	51.6	341	2	S43352	platelet-activatin
107	16	51.6	348	2	B75445	conserved hypotet
108	16	51.6	355	2	A34597	rhomboid protei
109	16	51.6	362	2	T36079	hypothetical prote
110	16	51.6	377	2	T05354	hypothetical prote
111	16	51.6	379	2	S14885	hypothetical prote
112	16	51.6	389	2	T29488	hypothetical prote
113	16	51.6	392	2	I54491	cell surface antig
114	16	51.6	416	1	JN0006	neve growth facto
115	16	51.6	418	2	T16713	hypothetical prote
116	16	51.6	429	2	I48361	gene G8G-154 prote
117	16	51.6	433	2	F86163	hypothetical prote
118	16	51.6	439	2	G97329	anaerobic dicarbox
119	16	51.6	468	2	T22397	hypothetical prote
120	16	51.6	480	2	T23608	hypothetical prote
121	16	51.6	483	2	T17346	hypothetical prote
122	16	51.6	491	2	T22844	hypothetical prote
123	16	51.6	492	2	S49147	BRF-2 protein - hu
124	16	51.6	496	2	S55665	conserved hypotet
125	16	51.6	521	2	F64522	conserved hypotet
126	16	51.6	523	2	B71985	hypothetical prote
127	16	51.6	525	2	G71317	hypothetical prote
128	16	51.6	528	2	D70968	hypothetical prote
129	16	51.6	543	2	F72071	conserved hypotet
130	16	51.6	543	2	C86551	CT384 hypothetical
131	16	51.6	545	2	H83342	periplasmic crehal
132	16	51.6	558	2	A36347	glypican 1 precurs
133	16	51.6	558	2	I56545	glypican precursor
134	16	51.6	561	2	A55335	rhoprty protein 2
135	16	51.6	569	2	T22703	hypothetical prote
136	16	51.6	583	2	UT0395	DNA-binding protei
137	16	51.6	596	2	S59944	hydroxymethylgluta
138	16	51.6	604	2	T04120	hypothetical prote
139	16	51.6	618	2	S33044	hydroxymethylgluta
140	16	51.6	625	2	S48941	regulatory protei
141	16	51.6	625	2	A48579	trophozoite surfac
142	16	51.6	667	2	T49326	hypothetical prote
143	16	51.6	687	2	C55926	DNA binding protei
144	16	51.6	692	2	A96590	hypothetical prote
145	16	51.6	700	2	A35502	major surface-labe
146	16	51.6	713	2	UC7518	subtilisin-like be
147	16	51.6	736	2	A36932	integrin beta-8 ch
148	16	51.6	768	2	B41029	iron(II) transport
149	16	51.6	773	1	A36932	iron(II) transport
150	16	51.6	773	2	C91160	ferrous iron trans
151	16	51.6	773	2	B86006	ferrous iron trans
152	16	51.6	797	2	A36811	hypothetical prote
153	16	51.6	829	1	PLMS	plasma( EC 3.4.21
154	16	51.6	829	2	T07406	probable protein k
155	16	51.6	864	2	S61148	SBE2 protei
156	16	51.6	882	2	E96931	hypothetical prote
157	16	51.6	925	2	T02811	DNA excision/repai
158	16	51.6	939	2	C81730	metalloprotease,
159	16	51.6	942	2	B72015	metalloprotease,
160	16	51.6	942	2	C86610	insulinase family/
161	16	51.6	953	2	I48078	CHOI antigen - Chi
162	16	51.6	962	2	JC5571	subtilisin-like pr
163	16	51.6	962	2	A33626	fibrinogen alpha c
164	16	51.6	969	1	A39490	subtilisin-like pr
165	16	51.6	975	2	UC5570	subtilisin-like pr
166	16	51.6	979	2	A89633	protein F13B6.6 [l
167	16	51.6	985	2	S67255	REVI protei
168	16	51.6	997	2	A87320	TonB-dependent rec
169	16	51.6	1011	2	S65668	preprote
170	16	51.6	1058	2	A53767	mucin MUC5B, trach
171	16	51.6	1058	2	T30178	mitotic checkpoint
172	16	51.6	1102	2	T31004	probable mitotic c
173	16	51.6	1111	2	T01239	hypothetical prote
174	16	51.6	1111	2	T26972	hypothetical prote
175	16	51.6	1151	2	H71347	hypothetical prote
176	16	51.6	1176	2	T49482	hypothetical prote
177	16	51.6	1274	2	T42017	cysteine rich prot
178	16	51.6	1609	2	S25345	probable membrane
179	16	51.6	1729	2	A49282	fusion protei
180	16	51.6	1899	2	T49273	hypothetical prote
181	16	51.6	2105	2	T18968	probable serine-ty
182	16	51.6	2233	2	T28669	surface protei
183	15	48.4	25	1	SMKR	metallothionein
184	15	48.4	25	1	S52359	hypothetical prote
185	15	48.4	43	2	D84641	hypothetical prote
186	15	48.4	44	2	T19161	hypothetical prote
187	15	48.4	46	1	VTVA1P	viscotoxin 1-PS -
188	15	48.4	46	1	VTVA1P	viscotoxin A2 - Bu
189	15	48.4	46	1	VTVA1P	viscotoxin B - Bur
190	15	48.4	52	2	UC1197	metallothionein II
191	15	48.4	59	1	B44203	ferredoxin [3Fe-4S
192	15	48.4	60	2	S44350	ferredoxin [4Fe-4S
193	15	48.4	61	2	SMHY1C	metallothionein I
194	15	48.4	61	2	I65234	testicular lutein
195	15	48.4	61	2	S00810	metallothionein IC
196	15	48.4	62	2	S15676	cholesterol gonadot
197	15	48.4	62	2	B55995	prostaglandin E2 r
198	15	48.4	67	2	T01958	hypothetical prote
199	15	48.4	77	2	I48725	Q300 protei
200	15	48.4	78	2	S25619	hypothetical prote
201	15	48.4	78	2	D72764	probable ferredoxi
202	15	48.4	87	2	UT0623	Killer toxin HYI
203	15	48.4	88	2	JQ0514	cysteine-rich 9K p
204	15	48.4	88	2	S12126	cysteine-rich oute
205	15	48.4	88	2	S12125	cysteine-rich oute
206	15	48.4	88	2	E71513	probable 9kDa-cy
207	15	48.4	88	2	D81671	cysteine rich oute
208	15	48.4	92	1	KVRB38	Ig kappa chain V r
209	15	48.4	93	2	T10379	Ig kappa chain V r
210	15	48.4	95	1	FEDH1	ferredoxin [2Fe-2S
211	15	48.4	95	1	FEDH2	ferredoxin [2Fe-2S
212	15	48.4	97	2	T03100	hypothetical prote
213	15	48.4	101	2	T48791	hypothetical prote
214	15	48.4	101	2	T51856	RIN6-H2 finger pro
215	15	48.4	103	2	H84599	hypothetical prote
216	15	48.4	104	1	KVRBXP	Ig kappa chain V r
217	15	48.4	104	2	S60405	hypothetical prote
218	15	48.4	107	1	KARB41	Ig kappa chain V r
219	15	48.4	107	2	A72782	hypothetical prote
220	15	48.4	108	1	KVRB12	Ig kappa chain V r
221	15	48.4	108	1	KVRB2K	Ig kappa chain V r
222	15	48.4	108	1	KVRB2K	Ig kappa chain V r
223	15	48.4	108	1	KVRB2K	Ig kappa chain V r
224	15	48.4	108	1	KVRB2K	Ig kappa chain V r
225	15	48.4	109	1	KVRB37	Ig kappa chain V r
226	15	48.4	109	1	PVSNBB	Ig kappa chain V r
227	15	48.4	109	1	D71138	Ig kappa chain V r
228	15	48.4	110	1	KVRB29	Ig kappa chain V r
229	15	48.4	110	1	KVRB54	Ig kappa chain V r
230	15	48.4	112	2	S21196	Ig kappa chain V r
231	15	48.4	112	2	A84605	hypothetical prote
232	15	48.4	113	1	WNVNEM	EcBRI-T medium (ET
233	15	48.4	113	2	PC1274	NS5 protei
234	15	48.4	113	2	PC1277	NS5 protei
235	15	48.4	113	2	PC1278	NS5 protei
236	15	48.4	113	2	PC1276	NS5 protei
237	15	48.4	113	2	PC1275	NS5 protei
238	15	48.4	114	2	S57692	probable membrane
239	15	48.4	117	1	K4RFP2	Ig kappa chain pre
240	15	48.4	117	1	BVBCAR	arsenical resistan
241	15	48.4	117	2	G91131	probable truncated
242	15	48.4	118	2	PN0141	lutropin beta chai
243	15	48.4	118	2	PN0139	lutropin beta chai
244	15	48.4	118	2	B32536	T-cell receptor al
245	15	48.4	119	2	A61465	lutropin beta chai
246	15	48.4	120	2	D82022	truncated pili
247	15	48.4	121	2	S58424	seminal plasma pro
248	15	48.4	121	2	D69047	ychn protei

249	15	48.4	122	2	A27635	Ig heavy chain pre	322	15	48.4	189	2	T19559	hypothetical prote
250	15	48.4	122	2	C64965	hypothetical prote	323	15	48.4	190	2	C69345	hypothetical prote
251	15	48.4	122	2	B85825	probable structura	324	15	48.4	191	2	AH3005	conserved hypotet
252	15	48.4	122	2	E90979	probable structura	325	15	48.4	191	2	S69735	hypothetical prote
253	15	48.4	125	2	S53562	probable membrane	326	15	48.4	193	2	E95402	hypothetical prote
254	15	48.4	125	2	B70981	hypothetical prote	327	15	48.4	194	2	UQ1560	hypothetical 20.6k
255	15	48.4	125	2	T01628	hypothetical prote	328	15	48.4	194	2	T41796	hypothetical 20.6k
256	15	48.4	127	2	T44119	hypothetical prote	329	15	48.4	195	2	T07735	nodulin-20a - soyb
257	15	48.4	127	2	B83104	hypothetical prote	330	15	48.4	196	2	PN0478	vancomycin resista
258	15	48.4	127	2	C84302	hypothetical prote	331	15	48.4	197	2	A81877	probable membrane
259	15	48.4	128	2	T18072	hypothetical prote	332	15	48.4	197	2	F72628	hypothetical prote
260	15	48.4	132	2	B24255	chorion class A pr	333	15	48.4	204	2	T28939	hypothetical prote
261	15	48.4	132	2	PQ0394	genome polypotein	334	15	48.4	205	2	S37804	hypothetical prote
262	15	48.4	132	2	PQ0396	genome polypotein	335	15	48.4	206	2	UC7311	20k proteain - soyb
263	15	48.4	134	2	D90229	hypothetical prote	336	15	48.4	206	2	T25139	hypothetical prote
264	15	48.4	135	2	S36197	hypothetical prote	337	15	48.4	208	2	S49196	Kunitz trypsin inh
265	15	48.4	138	2	S00512	lutropin beta chai	338	15	48.4	209	2	A83849	alpha-ribazole-5'-
266	15	48.4	138	2	A82022	truncated piliin NM	339	15	48.4	209	2	S27494	nods proteain - Bra
267	15	48.4	139	2	I52320	testicular luteint	340	15	48.4	209	2	E96517	hypothetical prote
268	15	48.4	139	2	S72900	permease c2168b -	341	15	48.4	209	2	T29735	hypothetical prote
269	15	48.4	140	2	H84502	hypothetical prote	342	15	48.4	211	2	C88103	proteain M10G11.15
270	15	48.4	141	1	UTBOB	lutropin beta chai	343	15	48.4	212	2	T29906	hypothetical prote
271	15	48.4	141	1	UTPCB	lutropin beta chai	344	15	48.4	213	2	B25750	nodulin-26b - soyb
272	15	48.4	141	1	UTRTB	lutropin beta chai	345	15	48.4	213	2	A25750	nodulin-27 - soyb
273	15	48.4	141	1	UTSHB	lutropin beta chai	346	15	48.4	213	2	A55583	pyrogutamy1-pepti
274	15	48.4	141	2	JC4527	luteinizing hormon	347	15	48.4	214	2	D71282	probable deda prot
275	15	48.4	141	2	F82021	truncated piliin NM	348	15	48.4	215	2	C98278	hypothetical prote
276	15	48.4	142	2	JC6305	neutrin precursor	349	15	48.4	215	2	B57055	melanocortin-4 rec
277	15	48.4	143	2	S52595	probable membrane	350	15	48.4	215	2	A26669	nodulin-22 precurs
278	15	48.4	144	2	T50622	hypothetical prote	351	15	48.4	215	2	T24710	hypothetical prote
279	15	48.4	145	2	T42280	hypothetical prote	352	15	48.4	216	1	F69686	orotate phosphorib
280	15	48.4	145	2	T33037	hypothetical prote	353	15	48.4	216	2	S54430	fimtrial proteain h
281	15	48.4	146	2	T01683	immobilization ant	354	15	48.4	217	2	S01358	salivary glue prot
282	15	48.4	150	2	T49563	hypothetical prote	355	15	48.4	220	2	S52005	nodulin-30 (Mpv30)
283	15	48.4	150	2	D87652	hypothetical prote	356	15	48.4	220	2	S42879	nodulin-30 - kidne
284	15	48.4	152	2	A32669	vasotocin 1 / neur	357	15	48.4	223	2	A23036	nodulin-23 - soybe
285	15	48.4	152	2	AG0060	probable exported	358	15	48.4	224	2	S07315	nodulin - soybean
286	15	48.4	156	2	H83361	Arsc proteain PA227	359	15	48.4	224	2	S56809	probable membrane
287	15	48.4	158	2	S35787	US9 proteain - bovi	360	15	48.4	225	1	WZBBF1	gene 58 proteain -
288	15	48.4	159	2	B29879	vasotocin / neurop	361	15	48.4	225	2	D88492	proteain T07E3.6 (I
289	15	48.4	160	2	T05385	hypothetical prote	362	15	48.4	225	2	D64598	hypothetical prote
290	15	48.4	161	1	S43893	H+-exporting ATPas	363	15	48.4	226	2	H96547	probable BHLH tran
291	15	48.4	161	2	S14480	arginine-vasotocin	364	15	48.4	228	2	B72655	hypothetical prote
292	15	48.4	161	2	T13688	destablase 2 homo	365	15	48.4	228	2	H71350	hypothetical prote
293	15	48.4	162	2	P97139	hypothetical prote	366	15	48.4	228	2	S26995	cellulose 1,4-beta
294	15	48.4	164	2	A41712	H+-exporting ATPas	367	15	48.4	229	2	S60454	glucose starvation
295	15	48.4	165	2	AH0215	4-hydroxyphenylace	368	15	48.4	230	2	S18989	fama proteain - Scr
296	15	48.4	165	2	T48806	hypothetical prote	369	15	48.4	231	2	T24669	hypothetical prote
297	15	48.4	169	1	KTHOB	choriogonadotropin	370	15	48.4	234	2	S76342	hypothetical prote
298	15	48.4	170	2	AE0630	4-hydroxyphenylace	371	15	48.4	235	2	F82754	cellulose 1,4-beta
299	15	48.4	170	2	A83135	4-hydroxyphenylace	372	15	48.4	236	2	S26993	regulatory proteain
300	15	48.4	170	2	T06056	hypothetical prote	373	15	48.4	237	1	B42371	riboflavin proteain
301	15	48.4	172	1	B69138	hypothetical prote	374	15	48.4	238	1	VOCH	riboflavin-binding
302	15	48.4	173	2	B39141	transcription repr	375	15	48.4	238	1	LNRTMA	mannose-binding 1e
303	15	48.4	174	2	C69404	hypothetical prote	376	15	48.4	238	2	A64156	Mg2+-transporting
304	15	48.4	174	2	T49180	hypothetical prote	377	15	48.4	238	2	C83194	conserved hypotet
305	15	48.4	174	2	B85977	hypothetical prote	378	15	48.4	239	2	T16640	hypothetical prote
306	15	48.4	176	2	E71371	hypothetical prote	379	15	48.4	239	1	JC1373	ribonuclease (EC 3
307	15	48.4	177	2	S25551	thamatin-like pro	380	15	48.4	241	2	T30761	hypothetical prote
308	15	48.4	177	2	T07843	probable lysine-ke	381	15	48.4	242	2	UC4360	riboflavin-binding
309	15	48.4	179	2	T19557	hypothetical prote	382	15	48.4	242	2	T29699	hypothetical prote
310	15	48.4	179	2	P97324	uncharacterized co	383	15	48.4	243	2	H69525	hydroxymethylpyrim
311	15	48.4	180	2	S15327	fintrial proteain M	384	15	48.4	244	2	T31838	hypothetical prote
312	15	48.4	180	2	C23308	nifo proteain - Rho	385	15	48.4	245	2	T24565	proteolipid protei
313	15	48.4	180	2	F88028	proteain F46F5.15 (	386	15	48.4	246	2	I51325	mannose-binding 1e
314	15	48.4	181	2	T04165	pathogenesis-relat	387	15	48.4	248	1	LNHUMC	mannose-binding 1e
315	15	48.4	182	2	T26109	hypothetical prote	388	15	48.4	248	2	S49323	chymotrypsin (EC 3
316	15	48.4	184	2	C72740	hypothetical prote	389	15	48.4	248	2	B72045	uridylylate kinase C
317	15	48.4	185	2	B69028	hypothetical prote	390	15	48.4	248	2	G86577	UMP kinase [import
318	15	48.4	185	2	B26669	probable phosphono	391	15	48.4	248	2	D96535	RING-H2 finger pro
319	15	48.4	186	1	DEPSNL	nodulin-20 precurs	392	15	48.4	249	2	T11837	hypothetical prote
320	15	48.4	187	2	T32826	amine dehydrogenas	393	15	48.4	250	2	T21785	hypothetical prote
321	15	48.4	189	2	S60587	non-structural prote	394	15	48.4	250	1	A49053	CD27 antigen precu

395	15	48.4	250	2	T31836	hypothetical prote
396	15	48.4	250	2	T31835	hypothetical prote
397	15	48.4	253	2	T49483	hypothetical prote
398	15	48.4	255	2	B69962	amino acid ABC tra
399	15	48.4	256	2	B33393	T-cell antigen 4-1
400	15	48.4	257	2	T28946	hypothetical prote
401	15	48.4	258	2	G96798	hypothetical prote
402	15	48.4	259	2	T52619	TINy-like protein
403	15	48.4	260	1	A46517	CD27 antigen precu
404	15	48.4	262	1	VG1HPC	E1 membrane glycop
405	15	48.4	262	2	S47427	E1 membrane glycop
406	15	48.4	263	2	T30083	hypothetical prote
407	15	48.4	266	2	P00393	B27e genome polyprote
408	15	48.4	266	2	T36857	B27e genome polyprote
409	15	48.4	268	2	S08229	chd protein - var
410	15	48.4	270	2	I64041	chd protein - Hae
411	15	48.4	272	2	T20991	hypothetical prote
412	15	48.4	273	2	E75476	hypothetical prote
413	15	48.4	275	2	F69229	hypothetical prote
414	15	48.4	276	2	A80860	hypothetical prote
415	15	48.4	277	2	S38503	beta-lactamase (EC
416	15	48.4	277	2	S38671	beta-lactamase (EC
417	15	48.4	277	2	I37552	OX40 homolog - hum
418	15	48.4	279	2	B96917	membrane permease,
419	15	48.4	280	2	T02090	xyloglucan endo-1,
420	15	48.4	281	2	E64216	hypothetical prote
421	15	48.4	281	2	T03509	probable arsenate
422	15	48.4	285	2	G69204	ribose-phosphate p
423	15	48.4	285	2	T21962	hypothetical prote
424	15	48.4	286	2	T02119	probable DnaU prot
425	15	48.4	289	1	MF1HPC	E1 membrane glycop
426	15	48.4	291	2	T28846	hypothetical prote
427	15	48.4	292	2	D69176	conserved hypotet
428	15	48.4	294	1	E46M53	tissue factor proc
429	15	48.4	294	2	E84640	CONSTANS-like B-bo
430	15	48.4	294	2	T23682	hypothetical prote
431	15	48.4	300	2	T05729	hypothetical prote
432	15	48.4	303	2	T46715	hypothetical prote
433	15	48.4	303	2	T04541	hypothetical prote
434	15	48.4	305	2	F82765	conserved hypotet
435	15	48.4	305	2	T25117	hypothetical prote
436	15	48.4	307	2	T08125	CONSTANS protein h
437	15	48.4	307	2	D70741	hypothetical prote
438	15	48.4	307	2	T18846	hypothetical prote
439	15	48.4	312	2	A85178	probable flavoprot
440	15	48.4	312	2	E90929	probable flavoprot
441	15	48.4	312	2	B64928	electron transfer
442	15	48.4	312	2	B69842	3-oxoacyl- acyl-ca
443	15	48.4	312	2	B82239	bioin synthesis p
444	15	48.4	312	2	T31834	hypothetical prote
445	15	48.4	312	2	T32379	hypothetical prote
446	15	48.4	314	2	I37383	FAS soluble protei
447	15	48.4	314	2	S55179	mRNA splice defec
448	15	48.4	314	2	G84336	hypothetical prote
449	15	48.4	317	2	S29204	melanotropin recep
450	15	48.4	317	2	A64343	hypothetical prote
451	15	48.4	318	2	T28941	hypothetical prote
452	15	48.4	319	1	B42971	alcohol dehydrogen
453	15	48.4	320	2	S75911	hypothetical prote
454	15	48.4	320	2	A53119	cell adheson glyco
455	15	48.4	321	2	F84597	probable prolina-r
456	15	48.4	321	2	D81018	iron(III) ABC tran
457	15	48.4	321	2	B81962	probable membrane
458	15	48.4	321	2	T24773	hypothetical prote
459	15	48.4	321	2	D96035	probable transposa
460	15	48.4	321	2	E96025	probable transposa
461	15	48.4				



541	15	48.4	390	2	D97305	NADH-dependent but	614	15	48.4	485	2	T13479	hypothetical prote
542	15	48.4	390	2	B47013	butanol dehydrogen	615	15	48.4	487	2	T19875	hypothetical prote
543	15	48.4	392	2	C82628	hypothetical prote	616	15	48.4	490	2	T32003	hypothetical prote
544	15	48.4	392	2	T16776	hypothetical prote	617	15	48.4	493	2	S16931	chitinase (EC 3.2
545	15	48.4	393	2	T16046	hypothetical prote	618	15	48.4	498	2	T46203	interleukin-14 pre
546	15	48.4	393	2	S48288	probable phosphopr	619	15	48.4	500	2	A71306	hypothetical prote
547	15	48.4	394	2	H72505	hypothetical prote	620	15	48.4	500	2	C84771	probable glucosylt
548	15	48.4	394	2	F95973	probable sugar upt	621	15	48.4	501	2	S43536	T-cell surface gly
549	15	48.4	396	1	TRXXB4	alpha-lytic protei	622	15	48.4	501	2	JH0447	alpha-1A-adrenergi
550	15	48.4	396	1	DCBYD4	adenosylmethionine	623	15	48.4	501	2	E96758	hypothetical prote
551	15	48.4	396	1	E64987	bicyclomycin resis	624	15	48.4	504	1	B48939	cellulose 1,4-beta
552	15	48.4	396	2	B91013	bicyclomycin resis	625	15	48.4	504	1	S18450	variant surface gl
553	15	48.4	396	2	D85857	bicyclomycin resis	626	15	48.4	507	2	S14629	aldehyde dehydroge
554	15	48.4	398	2	F40511	hypothetical prote	627	15	48.4	510	2	S41943	cellulose 1,4-beta
555	15	48.4	398	2	S17428	interleukin-1 rece	628	15	48.4	510	2	A70580	probable UDP-N-ac
556	15	48.4	400	2	JC4591	alpha-1,3 fucosylt	629	15	48.4	511	2	S44716	cellulose 1,4-beta
557	15	48.4	400	2	T32705	hypothetical prote	630	15	48.4	512	2	A87270	hypothetical prote
558	15	48.4	405	2	B36340	alpha(1,3)-fucosyl	631	15	48.4	513	2	T37612	cytosol aminopepti
559	15	48.4	406	2	T01931	adenyllyl cyclase -	632	15	48.4	515	2	A40491	alpha-1-adrenergic
560	15	48.4	407	2	JE0113	zinc-finger protei	633	15	48.4	515	2	JC1525	alpha-1B-adrenergi
561	15	48.4	410	2	JC7757	function-unknown p	634	15	48.4	515	2	I39073	interferon alpha-b
562	15	48.4	411	2	T29475	hypothetical prote	635	15	48.4	516	2	S31164	cellulose 1,4-beta
563	15	48.4	414	2	H84596	hypothetical prote	636	15	48.4	516	2	JS0083	cellulose 1,4-beta
564	15	48.4	416	2	I45890	ferrochelatase (EC	637	15	48.4	517	2	A45121	alpha-1B adrenergi
565	15	48.4	417	2	B86919	probable seryl-tr	638	15	48.4	518	2	B86299	hypothetical prote
566	15	48.4	417	2	E30341	alpha-1-adrenergic	639	15	48.4	520	1	RD1NTS	dihydrofolate redu
567	15	48.4	417	2	T20199	hypothetical prote	640	15	48.4	520	1	RD1NTZ	dihydrofolate redu
568	15	48.4	418	2	E71716	proline/betaine tr	641	15	48.4	520	2	T20007	hypothetical prote
569	15	48.4	423	2	A36794	hypothetical prote	642	15	48.4	520	2	S14600	E2 glycoprotein pr
570	15	48.4	423	2	T19581	hypothetical prote	643	15	48.4	521	2	E82508	methyl-accepting c
571	15	48.4	425	2	D88115	protein F53C3.11 f	644	15	48.4	526	2	B81679	conserved hypotet
572	15	48.4	425	2	C89753	protein F11C7.3 f1	645	15	48.4	529	2	S62194	hypothetical prote
573	15	48.4	427	2	G88492	protein T07B3.4 f1	646	15	48.4	533	2	A45392	RNA-directed RNA p
574	15	48.4	429	1	T03607	probable carboxype	647	15	48.4	533	2	B56110	tyrosine phosphopr
575	15	48.4	431	2	T14414	S-locus-specific g	648	15	48.4	538	2	AB3632	D-ribulokinase (EC
576	15	48.4	432	2	A83616	dicarboxylic acid	649	15	48.4	539	2	D71522	hypothetical prote
577	15	48.4	432	2	T34154	hypothetical prote	650	15	48.4	540	2	S41942	cellulose 1,4-beta
578	15	48.4	433	2	T14329	dermal glycoprotei	651	15	48.4	540	2	B64829	membrane protein b
579	15	48.4	433	2	A57596	alpha-1,3-fucosylt	652	15	48.4	540	2	H90751	probable transport
580	15	48.4	434	2	G86163	hypothetical prote	653	15	48.4	540	2	F85615	probable transport
581	15	48.4	436	1	S49458	diphosphate-fructo	654	15	48.4	541	2	D84681	hypothetical prote
582	15	48.4	436	1	D69270	ferredoxin-nitrite	655	15	48.4	550	1	VG1HD6	E2 glycoprotein pr
583	15	48.4	439	1	ERADC2	early E1B 49K prot	656	15	48.4	552	1	S47034	cell fusion glycop
584	15	48.4	439	2	T15748	hypothetical prote	657	15	48.4	553	1	I46329	cell fusion glycop
585	15	48.4	445	2	A48881	serotonin receptor	658	15	48.4	553	1	VGNZGB	cell fusion glycop
586	15	48.4	445	2	T21744	hypothetical prote	659	15	48.4	553	1	A36830	cell fusion glycop
587	15	48.4	448	2	S36402	serotonin receptor	660	15	48.4	553	1	A46329	cell fusion glycop
588	15	48.4	448	1	A47519	serotonin receptor	661	15	48.4	553	1	B36830	cell fusion glycop
589	15	48.4	449	1	A42800	beta-alanine-pyruv	662	15	48.4	553	1	B46329	cell fusion glycop
590	15	48.4	450	2	E96695	hypothetical prote	663	15	48.4	553	1	D46329	cell fusion glycop
591	15	48.4	450	2	B36810	hypothetical prote	664	15	48.4	553	1	B46329	cell fusion glycop
592	15	48.4	451	2	T23731	hypothetical prote	665	15	48.4	553	1	G46329	cell fusion glycop
593	15	48.4	454	2	T26654	hypothetical prote	666	15	48.4	553	1	VGNZND	cell fusion glycop
594	15	48.4	456	2	T16796	hypothetical prote	667	15	48.4	553	1	VGNZNV	cell fusion glycop
595	15	48.4	458	2	B83341	L-serine dehydrata	668	15	48.4	553	1	VGNZTE	cell fusion glycop
596	15	48.4	458	2	T16123	hypothetical prote	669	15	48.4	553	1	VGNZUI	cell fusion glycop
597	15	48.4	460	2	C82399	conserved hypotet	670	15	48.4	553	2	S06345	cell fusion glycop
598	15	48.4	461	1	A35356	tumor necrosis fac	671	15	48.4	553	2	S38784	gene F protein - N
599	15	48.4	463	1	T43773	hypothetical prote	672	15	48.4	553	2	S40163	cell fusion protei
600	15	48.4	469	1	NM1V27	exo-alpha-sialidas	673	15	48.4	553	2	S38786	gene F protein - N
601	15	48.4	469	1	NM1V2	exo-alpha-sialidas	674	15	48.4	553	2	S23621	gene F protein - N
602	15	48.4	469	1	NM1V2	exo-alpha-sialidas	675	15	48.4	553	2	S23620	gene F protein - N
603	15	48.4	469	1	J01644	exo-alpha-sialidas	676	15	48.4	553	2	S38785	gene F protein - N
604	15	48.4	470	2	S78440	phosphoglucomutase	677	15	48.4	553	2	T15094	hypothetical prote
605	15	48.4	471	2	S36655	UDP rhamnose-antio	678	15	48.4	553	2	S23622	gene F protein - N
606	15	48.4	473	2	S60290	anthocyanin rhamo	679	15	48.4	554	2	B82934	hypothetical prote
607	15	48.4	473	2	AG0612	probable transport	680	15	48.4	554	2	A31300	perforin precursor
608	15	48.4	474	1	VYHUD	vitamin D-binding	681	15	48.4	554	2	A45818	cytochrome c
609	15	48.4	476	2	S09152	translation elonga	682	15	48.4	555	2	A37181	perforin 1 precurs
610	15	48.4	477	2	S71323	alpha-1A adrenergi	683	15	48.4	555	2	S63137	hypothetical prote
611	15	48.4	481	2	A89102	protein F25E5.4 f1	684	15	48.4	557	2	A46434	variant-specific g
612	15	48.4	483	2	B84454	hypothetical prote	685	15	48.4	560	2	A38731	alpha-1A adrenergi
613	15	48.4	483	2	T20895	hypothetical prote	686	15	48.4	563	2	C86397	protein TYN9.10 f1

687	15	48.4	567	2	D83400	conserved hypothet	760	15	48.4	737	2	AH1958	hypothetical prote
688	15	48.4	568	2	E96648	hypothetical prote	761	15	48.4	739	2	T12964	subtilisin homolog
689	15	48.4	571	2	S69210	protein kinase cak	762	15	48.4	740	2	T03847	fas-binding protei
690	15	48.4	572	2	I39369	alpha-1A-adrenergi	763	15	48.4	753	2	B36268	platelet glycoprot
691	15	48.4	573	2	T29880	hypothetical prote	764	15	48.4	753	2	T28787	hypothetical prote
692	15	48.4	572	2	T25397	hypothetical prote	765	15	48.4	761	2	JC5759	brain-specific ser
693	15	48.4	580	2	S72211	N-acetyl-beta-D-gl	766	15	48.4	766	2	S61424	inorganic diphosph
694	15	48.4	581	2	S17150	potassium channel	767	15	48.4	770	1	A38230	inorganic diphosph
695	15	48.4	580	2	D84523	probable auxin-reg	768	15	48.4	770	2	T00204	LDL receptor relat
696	15	48.4	590	2	A45283	interferon alpha/b	769	15	48.4	772	2	D96504	protein PgC16.25 (
697	15	48.4	590	2	A48461	ovarian abundant m	770	15	48.4	770	2	A60798	platelet glycoprot
698	15	48.4	597	2	J00107	hypothetical 66K p	771	15	48.4	778	2	T03156	ribonucleoside-dip
699	15	48.4	600	2	A47391	serum albumin prec	772	15	48.4	780	2	S43534	integrin beta3 - C
700	15	48.4	603	2	S28941	coagulation factor	773	15	48.4	781	2	UQ0317	hypothetical 82K p
701	15	48.4	605	1	ABPGS	serum albumin prec	774	15	48.4	784	2	A26547	platelet glycoprot
702	15	48.4	607	2	T34040	hypothetical prote	775	15	48.4	788	2	A26547	platelet glycoprot
703	15	48.4	608	1	ABONS1	serum albumin 1 pr	776	15	48.4	788	2	I77349	integrin beta-3 su
704	15	48.4	608	1	ABONS2	serum albumin 2 pr	777	15	48.4	788	2	I51530	hypothetical prote
705	15	48.4	608	1	ABRYS	serum albumin prec	778	15	48.4	792	2	T21276	probable lipase VC
706	15	48.4	608	2	S57632	serum albumin prec	779	15	48.4	796	2	H82406	peroxisome prolif
707	15	48.4	608	2	T19935	hypothetical prote	780	15	48.4	796	2	UC7355	hypothetical prote
708	15	48.4	609	1	ABHUS	serum albumin prec	781	15	48.4	796	2	T16424	F-spondin precurs
709	15	48.4	610	2	S35049	mucin JBR57 - huma	782	15	48.4	803	2	A47723	hypothetical prote
710	15	48.4	611	2	D96607	hypothetical prote	783	15	48.4	811	2	T25166	hypothetical prote
711	15	48.4	612	2	T05750	auxin-regulated pr	784	15	48.4	813	2	T21192	DNA polymerase III
712	15	48.4	612	2	S55084	probable membrane	785	15	48.4	815	2	T05555	hypothetical prote
713	15	48.4	613	2	A39402	potassium channel	786	15	48.4	818	2	T32154	hypothetical prote
714	15	48.4	613	2	T00853	hypothetical prote	787	15	48.4	819	2	C84615	hypothetical prote
715	15	48.4	614	2	E84006	formate dehydrogen	788	15	48.4	822	2	T01095	hypothetical prote
716	15	48.4	618	2	T00476	probable vacuolar	789	15	48.4	825	2	T29634	hypothetical prote
717	15	48.4	619	1	JH0776	hydrogenase (EC 1.	790	15	48.4	827	1	S28273	beta-galactosidase
718	15	48.4	619	2	T08613	hypothetical prote	791	15	48.4	828	2	S52393	transcription fact
719	15	48.4	620	2	T27008	hypothetical prote	792	15	48.4	833	2	T10695	ORF MSV089 probab
720	15	48.4	621	2	S72493	laccase (EC 1.10.3	793	15	48.4	836	2	T28250	hypothetical prote
721	15	48.4	621	2	JC7278	adaptor protein co	794	15	48.4	834	2	E96561	hypothetical prote
722	15	48.4	621	2	T22904	hypothetical prote	795	15	48.4	840	2	H86429	hypothetical prote
723	15	48.4	624	2	S22703	voltage-gated pota	796	15	48.4	842	2	T05400	hypothetical prote
724	15	48.4	624	2	G87619	sensor histidine k	797	15	48.4	851	2	S12159	env protein - huma
725	15	48.4	626	2	T04895	vacuolar sorting r	798	15	48.4	852	1	VCLJGG	genome polypept
726	15	48.4	627	2	H82573	exonuclease ABC s	799	15	48.4	852	1	VCLJGG	env polypept
727	15	48.4	627	2	B48442	membrane transport	800	15	48.4	852	2	T01364	homeodomain transc
728	15	48.4	627	4	A40201	arifact-warning s	801	15	48.4	858	1	VCLJG2	env polypept
729	15	48.4	629	2	S60385	probable membrane	802	15	48.4	859	1	VCLJST	env polypept
730	15	48.4	630	2	T43460	hypothetical prote	803	15	48.4	859	1	VCLJ22	env polypept
731	15	48.4	631	2	A54659	DNA repair protein	804	15	48.4	859	1	VCLJ22	env polypept
732	15	48.4	631	2	T31782	hypothetical prote	805	15	48.4	859	1	VCLJ22	env polypept
733	15	48.4	631	2	T29926	hypothetical prote	806	15	48.4	859	1	VCLJ22	env polypept
734	15	48.4	632	2	S42731	collagen alpha 1 c	807	15	48.4	859	1	VCLJ22	env polypept
735	15	48.4	632	2	A71259	probable dicarboxy	808	15	48.4	859	1	VCLJ22	env polypept
736	15	48.4	633	2	A36353	DNA repair protein	809	15	48.4	859	1	VCLJ22	env polypept
737	15	48.4	633	2	I64143	hypothetical prote	810	15	48.4	859	1	VCLJ22	env polypept
738	15	48.4	640	2	S69546	phosphoenolpyruvat	811	15	48.4	859	1	S24571	env polypept
739	15	48.4	653	2	H86373	protein T23E23.16	812	15	48.4	860	2	C96504	protein PgC16.23 (
740	15	48.4	654	2	B75587	probable N-glycosyl	813	15	48.4	860	2	T37768	probable vacuolar
741	15	48.4	663	2	B65025	hypothetical prote	814	15	48.4	861	2	A48825	Notch homolog Motc
742	15	48.4	663	2	A91048	probable 2-compone	815	15	48.4	862	2	S64821	probable membrane
743	15	48.4	664	2	T06598	dnak-type molecula	816	15	48.4	863	2	D88465	protein B0244.7 (i
744	15	48.4	665	1	A42792	succinate dehydrog	817	15	48.4	864	2	B86449	hypothetical prote
745	15	48.4	670	2	T32221	hypothetical prote	818	15	48.4	869	2	S53098	envelope polypept
746	15	48.4	678	2	H96552	hypothetical prote	819	15	48.4	870	2	A96637	hypothetical prote
747	15	48.4	679	2	B85892	hypothetical prote	820	15	48.4	884	2	T18649	hypothetical prote
748	15	48.4	682	2	S37159	NADPH-ferrihemopro	821	15	48.4	890	2	F84548	hypothetical prote
749	15	48.4	685	2	S49163	transferrin precur	822	15	48.4	892	2	A41697	nitrate assimilat
750	15	48.4	689	2	A43734	probable protein k	823	15	48.4	909	2	H87729	protein Y23H5A.7 (
751	15	48.4	703	2	B82148	ATP-dependent heli	824	15	48.4	915	2	T03589	probable aspartate
752	15	48.4	704	2	B84685	hypothetical prote	825	15	48.4	915	2	T16623	hypothetical prote
753	15	48.4	704	2	T38117	probable protein k	826	15	48.4	923	2	E83574	hypothetical prote
754	15	48.4	708	2	T19474	hypothetical prote	827	15	48.4	925	2	T37475	lipoprotein recept
755	15	48.4	711	2	T30107	hypothetical prote	828	15	48.4	929	2	T32492	hypothetical prote
756	15	48.4	712	1	VCLJ54	env polypept	829	15	48.4	931	2	T33744	hypothetical prote
757	15	48.4	720	2	A36526	choline O-acetyltr	830	15	48.4	942	2	C96574	hypothetical prote
758	15	48.4	728	2	T20561	hypothetical prote	831	15	48.4	946	2	T00024	ent-kaurene syntha
759	15	48.4	736	2	T12963	subtilisin homolog	832	15	48.4	959	2	E85276	hypothetical prote

833	15	48.4	960	2	T00808	hypothetical prote
834	15	48.4	964	1	VLJJC6	env polypepten pr
835	15	48.4	964	2	S13329	hypothetical prote
836	15	48.4	966	1	VLJJC	env polypepten pr
837	15	48.4	979	1	JC2349	protein-tyrosine-p
838	15	48.4	979	2	D96574	hypothetical prote
839	15	48.4	982	1	VLJLVS	env polypepten pr
840	15	48.4	982	2	VLJLVS	env polypepten pr
841	15	48.4	983	1	E45390	env polypepten pr
842	15	48.4	985	1	VLJLSP	insulin II gene en
843	15	48.4	989	2	T48845	env polypepten pr
844	15	48.4	990	1	G46335	env polypepten pr
845	15	48.4	993	2	A47500	Ig mu chain switch
846	15	48.4	993	2	S35633	DNA-binding protei
847	15	48.4	994	2	S18739	env protein - biml
848	15	48.4	996	2	T48721	PTP 35 protein - m
849	15	48.4	1011	2	T17430	tol protein - Neur
850	15	48.4	1014	2	P96501	hypothetical prote
851	15	48.4	1016	2	T05066	hypothetical prote
852	15	48.4	1022	2	T48358	hypothetical prote
853	15	48.4	1025	2	T45811	Ca2+-transporting
854	15	48.4	1025	2	T14453	Ca2+-transporting
855	15	48.4	1030	2	T00812	Ca2+-transporting
856	15	48.4	1032	2	T43257	beta-1,3 exoglucan
857	15	48.4	1042	2	T16169	hypothetical prote
858	15	48.4	1042	2	T29307	hypothetical prote
859	15	48.4	1044	2	E86613	ribonucleoside-dip
860	15	48.4	1044	2	A72010	ribonucleoside-dip
861	15	48.4	1047	2	F81728	probable ribonucle
862	15	48.4	1053	2	D71466	probable ribonucle
863	15	48.4	1054	2	T30177	semaphorin F precu
864	15	48.4	1074	2	JC5928	protein F1R22.14 l
865	15	48.4	1076	2	B96682	hypothetical prote
866	15	48.4	1082	2	T44177	hypothetical prote
867	15	48.4	1082	2	T43590	hypothetical prote
868	15	48.4	1086	2	T09325	probable capsid as
869	15	48.4	1086	2	T33693	hypothetical prote
870	15	48.4	1110	1	B42544	G2-G1 polypepten
871	15	48.4	1128	2	AD3008	peptide synthetase
872	15	48.4	1137	2	A98276	hypothetical prote
873	15	48.4	1147	2	S65440	nitric-oxide synth
874	15	48.4	1162	2	B97852	hypothetical prote
875	15	48.4	1162	2	I56985	kalinin B1 - mouse
876	15	48.4	1169	2	C71639	hypothetical prote
877	15	48.4	1170	2	A53612	laminin B1k chain
878	15	48.4	1172	1	TSHP2	thrombospondin 2 p
879	15	48.4	1172	2	A42587	thrombospondin 2 p
880	15	48.4	1184	1	A39804	thrombospondin pre
881	15	48.4	1184	2	T46039	hypothetical prote
882	15	48.4	1192	2	S69000	laminin gamma 2 ch
883	15	48.4	1205	2	T41987	hypothetical prote
884	15	48.4	1217	2	T00270	hypothetical prote
885	15	48.4	1230	2	T42735	TBP-interacting pr
886	15	48.4	1232	2	T43027	neural cell adhesi
887	15	48.4	1260	2	T01334	hypothetical prote
888	15	48.4	1292	2	T09229	galactose binding
889	15	48.4	1299	2	T43251	furin (EC 3.4.21.7
890	15	48.4	1315	2	T50262	probable nucleopor
891	15	48.4	1315	2	A28313	glued protein - fr
892	15	48.4	1353	2	JC4279	adenylate cyclase
893	15	48.4	1353	2	A03905	genome polypepten
894	15	48.4	1386	2	T00257	hypothetical prote
895	15	48.4	1391	2	S50608	hypothetical prote
896	15	48.4	1391	2	T20406	hypothetical prote
897	15	48.4	1414	1	S23809	collagen alpha 2(I
898	15	48.4	1414	2	T33236	hypothetical prote
899	15	48.4	1416	2	E88550	protein ZC94.1 [lm
900	15	48.4	1437	2	S07430	M polypepten prec
901	15	48.4	1437	2	P96783	unknown protein F2
902	15	48.4	1453	2	A36861	orf 1b protein - L
903	15	48.4	1500	2	T03824	probable immediate
904	15	48.4	1502	2	S45429	probable membrane
905	15	48.4	1513	2	T23681	hypothetical prote
906	15	48.4	1587	2	G86467	hypothetical prote
907	15	48.4	1599	2	T16210	hypothetical prote
908	15	48.4	1732	2	G84664	hypothetical prote
909	15	48.4	1753	2	T00350	hypothetical prote
910	15	48.4	1755	2	T51532	hypothetical prote
911	15	48.4	1765	2	T42388	sodium channel alp
912	15	48.4	1773	2	T05128	hypothetical prote
913	15	48.4	1817	2	T34249	hypothetical prote
914	15	48.4	1820	2	A55494	latent transformin
915	15	48.4	1896	2	B72175	D1SR protein - var
916	15	48.4	1897	2	T28621	hypothetical prote
917	15	48.4	1965	2	T33216	hypothetical prote
918	15	48.4	2103	1	JQ1621	genome polypepten
919	15	48.4	2115	2	S38480	nonstructural prot
920	15	48.4	2120	2	T30243	alpha tectorin - c
921	15	48.4	2133	2	T30637	hypothetical prote
922	15	48.4	2149	2	C96695	ribulose biphosph
923	15	48.4	2155	2	T30197	alpha tectorin - m
924	15	48.4	2180	2	T29764	hypothetical prote
925	15	48.4	2205	1	MMWVRN	nonstructural poly
926	15	48.4	2214	1	OZBYU2	pyrimidine synthe
927	15	48.4	2222	1	A36028	DNA-directed DNA p
928	15	48.4	2227	1	GNVTHB	genome polypepten
929	15	48.4	2227	1	GNVTHM	genome polypepten
930	15	48.4	2227	1	GNVTHR	genome polypepten
931	15	48.4	2227	1	GNVTHK	genome polypepten
932	15	48.4	2230	1	GNVTHA	genome polypepten
933	15	48.4	2254	2	T09053	low voltage-activa
934	15	48.4	2256	2	T00076	hypothetical prote
935	15	48.4	2276	2	T34022	zonadhesin - pig
936	15	48.4	2492	1	A44213	nonstructural poly
937	15	48.4	2492	1	C44213	nonstructural poly
938	15	48.4	2492	1	MMWVTD	alpha-51D immobi11
939	15	48.4	2533	2	T28675	alpha-51D immobi11
940	15	48.4	2533	2	T28674	alpha-51D immobi11
941	15	48.4	2543	2	T31687	sulfate antigen - p
942	15	48.4	2643	2	T29149	hypothetical prote
943	15	48.4	2731	2	VFIHJH	genome polypepten
944	15	48.4	2733	2	S15760	genome polypepten
945	15	48.4	2844	2	S28291	hypothetical prote
946	15	48.4	3005	2	T08841	polypepten - deou
947	15	48.4	3010	1	A45573	genome polypepten
948	15	48.4	3010	1	GNWVCJ	genome polypepten
949	15	48.4	3010	1	GNWVTC	genome polypepten
950	15	48.4	3010	1	GNWVTW	genome polypepten
951	15	48.4	3010	1	S18030	genome polypepten
952	15	48.4	3020	2	A43932	mucin 2 precursor,
953	15	48.4	3071	2	T45584	genome polypepten
954	15	48.4	3414	1	GNWVNE	genome polypepten
955	15	48.4	3507	2	T34513	hypothetical prote
956	15	48.4	3600	2	D86161	F1003.12 protein -
957	15	48.4	4116	2	T31379	calo protein - fru
958	15	48.4	4307	2	T20721	hypothetical prote
959	15	48.4	4600	2	T42737	gp330 protein prec
960	15	48.4	5138	2	B96695	hypothetical prote
961	15	48.4	5376	2	T44215	zonadhesin - mouse
962	15	48.4	6658	2	T13931	projectin - fruit
963	15	48.4	13288	2	T03099	mucin, submaxillar
964	15	48.4	13	2	B58533	CD61 homolog - cha
965	15	48.4	16	2	C59045	alpha-conotoxin Au
966	15	48.4	16	2	A59045	alpha-conotoxin Au
967	15	48.4	22	2	B55538	uvic protein - Pse
968	15	48.4	24	2	A27128	sperm histone - ra
969	15	48.4	26	1	SMNC	metallochionein -
970	15	48.4	26	1	S55029	CAP3 protein - ant
971	15	48.4	27	2	S55030	CAP3 protein - ant
972	15	48.4	34	2	B81213	hypothetical prote
973	15	48.4	37	2	B31252	metallochionein II
974	15	48.4	41	2	A59149	sigma-conotoxin GV
975	15	48.4	45	1	VSWTAL	purothionin A-1 -
976	15	48.4	46	1	VTVAA3	viscotoxin A3 - Eu
977	15	48.4	46	1	LORDAL	ligatoxin A - Arge
978	15	48.4	47	1	A24074	pyrularia thionin

979 14 45.2 47 2 S22828 protamine - killer  
980 14 45.2 47 2 I48949 cellular disintegr  
981 14 45.2 49 2 S00228 protamine - horse  
982 14 45.2 49 2 S02007 protamine I - rabb  
983 14 45.2 50 1 HSPG sperm histone - pi  
984 14 45.2 50 1 HSPR sperm histone - sh  
985 14 45.2 50 2 S22582 protamine 1 - sagu  
986 14 45.2 50 2 S21672 protamine 1 - pig  
987 14 45.2 50 2 T38209 probable metalloth  
988 14 45.2 51 1 HSBOS sperm histone P1 -  
989 14 45.2 51 1 HSMSS1 protamine - mouse  
990 14 45.2 51 2 S03997 protamine 1 - rat  
991 14 45.2 52 2 S65712 metallochionein 1  
992 14 45.2 53 2 S24596 metallochionein -  
993 14 45.2 53 2 S75905 hypothetical prote  
994 14 45.2 55 1 B7AG58 vifR7 protein prec  
995 14 45.2 56 1 N1BSSA sublin precursor  
996 14 45.2 56 2 T12783 sublinin 168 prec  
997 14 45.2 57 2 S12957 venom animal Kunlt  
998 14 45.2 58 1 SMD1S metallochionein 1  
999 14 45.2 58 2 G97810 hypothetical prote  
1000 14 45.2 59 1 FEDV1V ferredoxin [4Fe-4S

## ALIGNMENTS

## RESULT 1

T28945 hypothetical protein F07C4.2 - Caenorhabditis elegans

C:/Species: Caenorhabditis elegans

C:/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C:/Accession: T28945

R:/Miller, N.; Steellies, L.

submitted to the EMBL Data Library, January 1997

A:/Description: The sequence of C. elegans cosmid F07C4.

A:/Reference number: Z20546

A:/Accession: T28945

A:/Status: preliminary; translated from GB/EMBL/DBJ

A:/Molecule type: DNA

A:/Residues: 1-155 &lt;ML&gt;

A:/Cross-references: UNIPROT:P91214; UNIPARC:UPI000007DB2D; EMBL:U80023; P1DN:AA048012.1;

A:/Experimental source: strain Bristol N2; clone F07C4

C:/Genetics:

A:/Gene: CESP.F07C4.2

A:/Map position: 5

A:/Intons: 116/1

C:/Superfamily: tetranectin; C-type lectin homology

Query Match 58.1%; Score 18; DB 2; Length 155;

Best Local Similarity 22.2%; Pred. No. 17;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12

DB 92 CTTASSSSC 100

## RESULT 2

T45962 hypothetical protein F7J8.200 - Arabidopsis thaliana

C:/Species: Arabidopsis thaliana (mouse-ear cress)

C:/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004

C:/Accession: T45962

R:/Bevan, M.; Zimmermann, W.; Gruenisen, A.; Mambuti, R.; Bancroft, I.; Mewes, H.W.; Lem

submitted to the Protein Sequence Database, January 2000

A:/Reference number: Z23018

A:/Accession: T45962

A:/Status: preliminary

A:/Molecule type: DNA

A:/Residues: 1-470 &lt;BBV&gt;

A:/Cross-references: UNIPROT:Q9LFB4; UNIPARC:UPI00000A7D3B; EMBL:AL137189

A:/Experimental source: cultivar Columbia; BAC clone F7J8

C:/Genetics:  
A:/Map position: 5  
A:/Intons: 116/2; 138/3; 162/2; 201/3; 226/2; 250/1; 275/1; 288/2; 307/2; 379/3  
A:/Note: F7J8.200

Query Match 58.1%; Score 18; DB 2; Length 470;

Best Local Similarity 22.2%; Pred. No. 23;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12

DB 23 CSSSSATSC 31

## RESULT 3

A34598 ecdysone-induced protein E75A - fruit fly (Drosophila melanogaster)

C:/Species: Drosophila melanogaster

C:/Date: 29-Jun-1990 #sequence\_revision 29-Jun-1990 #text\_change 31-Dec-2004

C:/Accession: A34598

R:/Segraves, W.A.; Hogness, D.S.

A:/Title: The E75 ecdysone-inducible gene responsible for the 75B early puff in Drosophila

A:/Reference number: A34598; PMID:90249727; PMID:2110921

A:/Accession: A34598

A:/Status: preliminary

A:/Molecule type: DNA

A:/Residues: 1-1237 &lt;SEG&gt;

A:/Cross-references: UNIPROT:P17671; UNIPARC:UPI0000129B00; GB:X51548; NID:G7910; P1DN:CA

C:/Genetics:

A:/Gene: FlyBase:Rip75B

A:/Cross-references: FlyBase:FBgn0000568

C:/Keywords: alternative splicing; DNA binding; nucleus; transcription regulation; zinc f

F:/243-520/Domain: exBA transforming protein homology &lt;ERBA&gt;

Query Match

Best Local Similarity 22.2%; Score 18; DB 2; Length 1237;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12

DB 1070 CSSSSASSC 1078

## RESULT 4

B34598

ecdysone-induced protein E75B - fruit fly (Drosophila melanogaster)

C:/Species: Drosophila melanogaster

C:/Date: 29-Jun-1990 #sequence\_revision 29-Jun-1990 #text\_change 31-Dec-2004

C:/Accession: B34598

R:/Segraves, W.A.; Hogness, D.S.

A:/Title: The E75 ecdysone-inducible gene responsible for the 75B early puff in Drosophila

A:/Reference number: A34598; PMID:90249727; PMID:2110921

A:/Accession: B34598

A:/Status: preliminary

A:/Molecule type: DNA

A:/Residues: 1-1394 &lt;SEG&gt;

A:/Cross-references: UNIPROT:P17672; UNIPARC:UPI0000129B01; GB:X51549; NID:G7912; P1DN:CA

C:/Genetics:

A:/Gene: FlyBase:Rip75B

A:/Cross-references: FlyBase:FBgn0000568

C:/Keywords: alternative splicing; DNA binding; nucleus; transcription regulation; zinc

F:/401-677/Domain: exBA transforming protein homology &lt;ERBA&gt;

Query Match 58.1%; Score 18; DB 2; Length 1394;

Best Local Similarity 22.2%; Pred. No. 31;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12

DB 1227 CSSSSASSC 1235

## RESULT 5

S05979  
steroid hormone receptor homolog - fruit fly (Drosophila melanogaster)  
N/Alternate names: puff 75B protein  
C/Species: Drosophila melanogaster  
C/Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 05-Oct-2004  
C/Accession: S05979  
R/Feigl, G.; Gram, M.; Pongs, O.  
Nucleic Acids Res. 17, 7167-7178, 1989  
A/Title: A member of the steroid hormone receptor gene family is expressed in the 20-OH-  
A/Reference number: S05979; MUID:90016778; PMID:2508058  
A/Accession: S05979  
A/Molecule type: mRNA  
A/Residues: 1-1443 <PEI>  
A/Cross-references: UNIPROT:P13055; UNIPARC:UPI0000129B02; EMBL:X15586; NID:97516; PIDN:  
C/Genetics:  
A/Genes: FlyBase:Fly75B  
A/Cross-references: FlyBase:Fbgn0000568  
A/Map position: 3 75B  
C/Keywords: alternative splicing; DNA binding; steroid hormone receptor; transcription  
F/452-727/Domain: e2bA transforming protein homology <ERBA>

Query Match 58.1%; Score 18; DB 2; Length 1443;  
Best Local Similarity 22.2%; Pred. No. 32;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 1276 CXXXXXXC 1284

## RESULT 6

TJ33355  
hypothetical protein F16G10.4 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: TJ33355  
R/Gattung, S.; Scheet, P.  
submitted to the EMBL Data Library, July 1998  
A/Description: The sequence of C. elegans cosmid F16G10.  
A/Reference number: Z21329  
A/Accession: TJ33355  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-139 <GAT>  
A/Cross-references: UNIPROT:O76588; UNIPARC:UPI000007891D; EMBL:AF077537; PIDN:AAC26272.  
A/Experimental source: strain Bristol N2; clone F16G10  
C/Genetics:  
A/Genes: CRSP:F16G10.4  
A/Map position: 2  
A/Functions: 56/3; 69/1; 89/1; 133/3  
C/Superfamily: Caenorhabditis elegans hypothetical protein F16G10.4

Query Match 54.8%; Score 17; DB 2; Length 199;  
Best Local Similarity 22.2%; Pred. No. 51;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 72 CEATSTATC 80

## RESULT 7

S63342  
hypothetical protein YNR014w - yeast (Saccharomyces cerevisiae)  
N/Alternate names: hypothetical protein N2060  
C/Species: Saccharomyces cerevisiae  
C/Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 09-Jul-2004  
C/Accession: S63342; S63344; S63340  
R/Maurer, C.T.C.; Urbanus, J.H.M.; Planta, R.J.  
submitted to the Protein Sequence Database, April 1996  
A/Reference number: S63346

A/Accession: S63342  
A/Molecule type: DNA  
A/Residues: 1-212 <MAU>

A/Cross-references: UNIPROT:P53719; UNIPARC:UPI000013BAB3; EMBL:Z71629; NID:G1302493; P  
A/Experimental source: strain S288C  
R/Duesterhoeft, A.; Floeth, M.; Fritz, C.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D.  
submitted to the Protein Sequence Database, April 1996  
A/Reference number: S62944  
A/Accession: S63344  
A/Molecule type: DNA  
A/Residues: 1-212 <DUE>  
A/Cross-references: UNIPARC:UPI000013BAB3; EMBL:Z71629; NID:G1302493; PID:e239786; PID:  
A/Experimental source: strain S288C  
R/Aert, R.; Verhaesselt, P.; Voet, M.; Volckaert, G.  
submitted to the Protein Sequence Database, April 1996  
A/Reference number: S62910  
A/Accession: S63340  
A/Molecule type: DNA  
A/Residues: 1-26 <AER>  
A/Cross-references: UNIPARC:UPI0000179C25; EMBL:Z71629; MIPS:YNR014w  
C/Genetics:  
A/Cross-references: SGD:S0005297  
A/Map position: 14R

Query Match 54.8%; Score 17; DB 2; Length 212;  
Best Local Similarity 22.2%; Pred. No. 52;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 40 CXXXXXXC 48

## RESULT 8

C70812  
probable lpqQ protein - Mycobacterium tuberculosis (strain H37RV)  
C/Species: Mycobacterium tuberculosis  
C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C/Accession: C70812  
R/Cole, S.T.; Broch, R.; Parham, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,  
Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Ruter, S.; Steger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.  
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A/Reference number: A70500; MUID:98295987; PMID:9634230  
A/Accession: C70812  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-214 <COL>  
A/Cross-references: UNIPROT:O53846; UNIPARC:UPI0000165256; GB:AL022004; GB:AL123456; NIT  
A/Experimental source: strain H37RV  
C/Genetics:  
A/Genes: lpqQ

Query Match 54.8%; Score 17; DB 2; Length 214;  
Best Local Similarity 22.2%; Pred. No. 52;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 15 CAIAITTAC 23

## RESULT 9

H72532  
hypothetical protein ABE2236 - Aeropyrum pernix (strain K1)  
C/Species: Aeropyrum pernix  
C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C/Accession: H72532  
R/Kawababayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai  
awa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999  
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: H72532  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-244 <KAW>  
A:Cross-references: UNIPROT:Q9Y9Q2; UNIPARC:UPI000005E27C; DDBJ:AP000063; NID:G5105654;  
A:Experimental source: strain KI  
C:Genetics:  
A:Gene: APE2236  
C:Superfamily: Aeropyrum pernix hypothetical protein APE2236

Query Match 54.8%; Score 17; DB 2; Length 244;  
Best Local Similarity 22.2%; Pred. No. 54;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 107 CSSSSAALC 115

RESULT 10  
T40415  
hypothetical protein SPBC4.01 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T40415  
R:McDougall, R.C.; Rajadream, M.A.; Barrell, B.G.; Brown, S.; Harris, D.  
submitted to the EMBL Data Library, October 1999  
A:Reference number: Z21927  
A:Accession: T40415  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-248 <MCD>  
A:Cross-references: UNIPROT:Q9USS9; UNIPARC:UPI000006AF49; EMBL:AL121863; PIDN:CAB58400;  
A:Experimental source: strain 972h-; cosmid c4  
C:Genetics:  
A:Gene: SPDB:SPBC4.01  
A:Map position: 2  
A:introns: 230/2

Query Match 54.8%; Score 17; DB 2; Length 248;  
Best Local Similarity 22.2%; Pred. No. 54;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 146 CLTTSFAC 154

RESULT 11  
H71353  
conserved hypothetical protein TP0182 - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: H71353  
R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
reese, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uteerback, T.; MCD  
science 281, 375-388, 1998  
A>Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; MUID:9832770; PMID:9665876  
A:Accession: H71353  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-259 <COL>  
A:Cross-references: UNIPROT:O83212; UNIPARC:UPI00000CA51; GB:AE01202; GB:AE000520; NID  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0182  
C:Superfamily: RNA-binding protein, YrdC type

Query Match 54.8%; Score 17; DB 2; Length 259;  
Best Local Similarity 22.2%; Pred. No. 55;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 31 CASASLSSC 39

RESULT 12  
AC0661  
probable lipoprotein SRY1394 [imported] - Salmonella enterica subsp. enterica serovar Typh  
C:Species: Salmonella enterica subsp. enterica serovar Typh  
A>Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AC0661  
R:Fairhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.,  
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AC0661  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-281 <PAR>  
A:Cross-references: UNIPARC:UPI0000059FOE; GB:AL513382; PIDN:CAD01660.1; PID:g16502512;  
C:Genetics:  
A:Gene: SRY1394

Query Match 54.8%; Score 17; DB 2; Length 281;  
Best Local Similarity 22.2%; Pred. No. 56;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 208 CATAAABC 216

RESULT 13  
T48975  
xyloglucan endo-transglycosylase - Arabidopsis thaliana  
N:Alternate names: protein F14D17.60  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
C:Accession: T48975  
R:Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.,  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z25008  
A:Accession: T48975  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-293 <JOR>  
A:Cross-references: UNIPROT:P93046; UNIPARC:UPI00000487CE; EMBL:AL353992; GSPDB:GN00061  
A:Experimental source: cultivar Columbia; BAC clone F14D17  
C:Genetics:  
A:Gene: ATSP:F14D17.60  
A:Map position: 3  
A:introns: 68/1; 101/3; 174/2  
C:Superfamily: endoxylglucan transferase

Query Match 54.8%; Score 17; DB 2; Length 293;  
Best Local Similarity 22.2%; Pred. No. 57;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 238 CTADSSSC 246

RESULT 14  
T01948

hypothetical protein F1104.13 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 09-Jul-2004  
C/Accession: T01948  
R/Abu-Threiden, J.; Stoneking, T.; Langston, Y.; Trevasakis, E.  
submitted to the EMBL Data Library, October 1998  
A/Description: The sequence of A. thaliana F1104.  
A/Reference number: Z14466  
A/Accession: T01948  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-298 <ABU>  
A/Cross-references: UNIPROT:O82599; UNIPARC:UP100000A5CDC; EMBL:AF096370; NID:g3695372;  
C/Genetics:  
A/Map position: 4  
A/Intons: 158/3  
A/Note: F1104.13  
C/Superfamily: Arabidopsis thaliana hypothetical protein F21F14.120

Query Match 54.8%; Score 17; DB 2; Length 298;  
Best Local Similarity 22.2%; Pred. No. 57;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 200 CXXXXXXC 208

RESULT 15  
A85019  
probable bHLH DNA-binding protein [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: A85019  
R/Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
Nature 402, 769-777, 1999  
A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A/Reference number: A85001; MUID:20083488; PMID:10617198  
A/Accession: A85019  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-315 <STO>  
A/Cross-references: UNIPROT:Q9M128; UNIPARC:UP1000009FFA; GB:NC\_001268; NID:g7268189; F  
C/Genetics:  
A/Gene: AT4G01460  
A/Map position: 4  
C/Superfamily: Arabidopsis thaliana hypothetical protein F21F14.120

Query Match 54.8%; Score 17; DB 2; Length 315;  
Best Local Similarity 22.2%; Pred. No. 58;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 188 CXXXXXXC 196

RESULT 16  
AB1224  
cobalamin biosynthesis protein cbid homolog cbid [imported] - Listeria monocytogenes (st  
C/Species: Listeria monocytogenes  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AB1224  
R/Glaeser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A/Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; M  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend  
A/Title: Comparative genomics of Listeria species.  
A/Reference number: AB1077; MUID:21537279; PMID:11679669  
A/Accession: AB1224

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-373 <GLA>  
A/Cross-references: UNIPROT:Q8Y757; UNIPARC:UP10000054CCA; GB:NC\_003210; PIDN:CAC99272.  
C/Genetics:  
A/Genetic source: strain EGD-e  
C/Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 54.8%; Score 17; DB 2; Length 373;  
Best Local Similarity 22.2%; Pred. No. 61;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 21 CAAAAKAC 29

RESULT 17  
AD1577  
cobalamin biosynthesis protein cbid homolog cbid [imported] - Listeria innocua (strain  
C/Species: Listeria innocua  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AD1577  
R/Glaeser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihi, H  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A/Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maltournam, A.; M  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend  
A/Title: Comparative genomics of Listeria species.  
A/Reference number: AB1077; MUID:21537279; PMID:11679669  
A/Accession: AD1577  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-373 <GLA>  
A/Cross-references: UNIPROT:Q92C44; UNIPARC:UP100000CC44; GB:AL592022; PIDN:CAC9388.1  
A/Experimental source: strain Clp11262  
C/Genetics:  
A/Gene: cbid  
C/Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 54.8%; Score 17; DB 2; Length 373;  
Best Local Similarity 22.2%; Pred. No. 61;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 21 CAAAAKAC 29

RESULT 18  
T43352  
nuclear receptor NHR-13 - Caenorhabditis elegans (fragment)  
C/Species: Caenorhabditis elegans  
C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C/Accession: T43352  
R/Sluder, A.E.; Mathews, S.W.; Hough, D.; Yin, V.P.; Maina, C.V.  
Genome Res. 9, 103-120, 1999  
A/Title: The nuclear receptor superfamily has undergone extensive proliferation and div  
A/Reference number: Z22443; MUID:99148134; PMID:10022975  
A/Accession: T43352  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-415 <SLU>  
A/Cross-references: UNIPROT:Q9YXB8; UNIPARC:UP10000075586; EMBL:AF083228; NID:g4139081;  
C/Genetics:  
A/Gene: nhr-13  
A/Map position: 5

Query Match 54.8%; Score 17; DB 2; Length 415;  
Best Local Similarity 22.2%; Pred. No. 63;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 12 CSSSSNSSC 20

## RESULT 19

T33024 hypothetical protein K07H8.2 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T33024

R:Fullon, B.; Hawkins, J.; Gattung, S.; Woldmann, P.; Elliott, G.

submitted to the EMBL Data Library, February 1998

A:Description: The sequence of *C. elegans* cosmid K07H8.

A:Reference number: Z21264

A:Accession: T33024

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-507 <PUL>

A:Cross-references: UNIPROT:Q45182; UNIPARC:UPI0000080E7E; EMBL:AF047659; PIDN:AA04427.

A:Experimental source: strain Bristol N2; clone K07H8

C:Genetics:

A:Gene: CESP:K07H8.2

A:Map position: 4

A:Introns: 20/3; 55/2; 93/3; 152/3; 447/3

Query Match 54.8%; Score 17; DB 2; Length 507;  
Best Local Similarity 22.2%; Pred. No. 66;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 210 CASSIATAC 218

## RESULT 20

T48004 multifunctional aminoacyl-tRNA ligase-like protein - *Arabidopsis thaliana*

N:Alternate names: protein T17J13.80

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C:Accession: T48004

R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Mayer,

submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24482

A:Accession: T48004

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-530 <PIE>

A:Cross-references: UNIPROT:Q9M1R2; UNIPARC:UPI000009CD43; EMBL:AL138651

A:Experimental source: cultivar Columbia; BAC clone T17J13

C:Genetics:

A:Map position: 3

A:Introns: 64/3; 94/3; 133/3; 196/2; 218/3; 259/3; 310/3; 405/3; 478/3; 510/1

A:Note: T17J13.80

C:Superfamily: proline-tRNA ligase pros

Query Match 54.8%; Score 17; DB 2; Length 530;  
Best Local Similarity 22.2%; Pred. No. 67;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 354 CTATAGALC 362

## RESULT 21

A42143 skn-1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A42143; T29535

R:Bowerman, B.; Eaton, B.A.; Priess, J.R.

Cell 68, 1061-1075, 1992

A>Title: skn-1, a maternally expressed gene required to specify the fate of ventral blas

A:Reference number: A42143; MUID:92191285; PMID:1547503

A:Accession: A42143

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-533 <BOV>

A:Cross-references: UNIPROT:Q6MPW3; UNIPARC:UPI0000081925

A:Note: sequence extracted from NCBI backbone (NCBIN:88973, NCBI:P:88974)

R:Pauley, A.; Gattung, S.

submitted to the EMBL Data Library, December 1995

A:Description: The sequence of *C. elegans* cosmid T19E7.

A:Reference number: Z20637

A:Accession: T29535

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-386 <PAU>

A:Cross-references: UNIPARC:UPI000017A521; EMBL:U42843; PIDN:AAA83594.1

C:Genetics:

A:Gene: CESP:skn-1

A:Introns: 141/3; 249/1; 300/1

C:Keywords: nucleus

F:495-532/Domain: fos/jun DNA-binding domain homology <PUD>

Query Match 54.8%; Score 17; DB 2; Length 533;  
Best Local Similarity 22.2%; Pred. No. 67;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 384 CTTDSSTC 392

## RESULT 22

T07739

probable ferrochelatase (EC 4.99.1.1) - potato

C:Species: *Solanum tuberosum* (potato)

C>Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004

C:Accession: T07739

R:Johnston, D.J.

submitted to the EMBL Data Library, April 1998

A:Reference number: Z15932

A:Accession: T07739

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-543 <JOH>

A:Cross-references: UNIPROT:Q64391; UNIPARC:UPI00000A543B; EMBL:AJ005802; PIDN:CAA06705

A:Experimental source: cv. Bintje

C:Function:

A:Description: catalyzes the insertion of iron into protoporphyrin to produce heme

A:Pathway: heme biosynthesis

C:Superfamily: ferrochelatase

C:Keywords: iron; lyase; mitochondrial inner membrane

Query Match 54.8%; Score 17; DB 2; Length 543;  
Best Local Similarity 22.2%; Pred. No. 67;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 31 CTSSFAASC 39

## RESULT 23

THHUB

thrombomodulin precursor [validated] - human

C:Species: *Homo sapiens* (man)

C>Date: 31-Dec-1988 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004

C:Accession: A41442; A28307; A29680; A27073; JX0264; S39554

R:Shitai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maruyama,

J. Biochem. 103, 281-285, 1988

A>Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed acti



A:Reference number: A41442; MUID:88227901; PMID:2836377  
 A:Accession: A41442  
 A:Molecule type: DNA  
 A:Residues: 1-575 <SHI>  
 A:Cross-references: UNIPROT:P07204; UNIPARC:UPI00000498EB; DDBJ:D00210; NID:g220126; PIR:J12600; R:Jackman, R.W.; Beeler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987  
 A:Title: Human thrombomodulin gene is intron depleted; nucleic acid sequences of the cDN  
 A:Reference number: A28307; MUID:87317665; PMID:2819876  
 A:Accession: A28307  
 A:Molecule type: DNA; mRNA  
 A:Residues: 1-472, 'A', 474-575 <JAC>  
 A:Cross-references: UNIPARC:UPI00000002BD; GB:J02973; NID:g339658; PIDN:AAA61175.1; PID:R:Suzuki, K.; Kusumoto, H.; Deysahtiki, Y.; Nishioka, J.; Matsuyama, I.; Zushi, M.; Kawahara  
 EMBO J. 6, 1891-1897, 1987  
 A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on endoth  
 A:Reference number: A29680; MUID:88004395; PMID:2820710  
 A:Accession: A29680  
 A:Molecule type: mRNA  
 A:Residues: 1-575 <SUZ>  
 A:Cross-references: UNIPARC:UPI00000498PB; GB:X05495; NID:g37123; PIDN:CAA29045.1; PID:G  
 A:Experimental source: lung endothelium  
 A:Note: Part of this sequence, including the amino end of the mature protein, were deter  
 R:Wen, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.  
 Biochemistry 26, 4350-4357, 1987  
 A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization of the  
 A:Reference number: A27073; MUID:88024950; PMID:2822087  
 A:Accession: A27073  
 A:Molecule type: mRNA  
 A:Residues: 1-472, 'A', 474-575 <WEN>  
 A:Cross-references: UNIPARC:UPI00000002BD; GB:M16552; NID:g339656; PIDN:AAB59508.1; PID:G  
 A:Experimental source: placenta  
 A:Note: parts of this sequence were determined by protein sequencing  
 R:Tamamoto, S.; Mizoguchi, T.; Tamaki, T.; Ohkuchi, M.; Kimura, S.; Aoki, N.  
 J. Biochem. 113, 433-440, 1993  
 A:Title: Urinary thrombomodulin, its isolation and characterization.  
 A:Reference number: JX0264; MUID:93293792; PMID:8390446  
 A:Accession: JX0264  
 A:Molecule type: protein; mRNA  
 A:Residues: 19-472, 'A', 474-486 <YAM>  
 A:Cross-references: UNIPARC:UPI0000173341  
 A:Experimental source: urine  
 A:Note: the urinary form appears to be identical with that circulating in plasma  
 R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell, B.W.  
 Biochem. J. 295, 131-140, 1993  
 A:Title: Identification of the predominant glycosaminoglycan-attachment site in soluble  
 A:Reference number: S38954; MUID:94029900; PMID:8216207  
 A:Accession: S38954  
 A:Molecule type: protein  
 A:Residues: 475-491, 'X', 493-494 <GER>  
 A:Cross-references: UNIPARC:UPI0000173342  
 A:Note: the residue designated 'X' was determined to be a Ser with covalently bound chon  
 R:Meininger, D.P.; Komives, E.A.  
 submitted to the Brookhaven Protein Data Bank, September 1995  
 A:Reference number: A67369; PDB:1ZAO  
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
 R:Fullinaky, A.; Mathews, I.I.  
 submitted to the Brookhaven Protein Data Bank, August 1994  
 A:Reference number: A52804; PDB:1HLT  
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442  
 R:Hrabal, R.; Komives, E.A.; Ni, F.  
 submitted to the Brookhaven Protein Data Bank, November 1995  
 A:Reference number: A65583; PDB:1FGD  
 A:Contents: annotation; conformation by (1)H-NMR, residues 427-444  
 R:Hrabal, R.; Komives, E.A.; Ni, F.  
 Protein Sci. 5, 195-203, 1996  
 A:Title: Structural resiliency of an EGF-like subdomain bound to its target protein, the  
 A:Reference number: A58595; MUID:96276211; PMID:8745396  
 A:Contents: annotation; conformation by (1)H-NMR  
 C:Genetics:  
 A:Gene: GDB:THBD  
 A:Cross-references: GDB:119613; OMIM:188040

A:Map position: 20p11.2-20p11.2  
 A:introns: #status absent  
 C:Complex: homodimer, urinary form  
 C:Function:  
 A:Description: inhibits thrombin activation of fibrinogen; cofactor for thrombin activa  
 A:Pathway: blood coagulation/moderation  
 A:Note: the membrane-bound form is located on the endothelium luminal surface of arteri  
 A:Note: thrombin complexed with the membrane-bound form is subject to endocytosis  
 C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
 C:Keywords: anticoagulant; beta-hydroxyasparagine; beta-hydroxyaspartic acid; blood coag  
 e protein  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>  
 F:19-513/Domain: extracellular #status predicted <EXT>  
 F:19-486/Product: thrombomodulin, urinary form #status experimental <MAU>  
 F:24-167/Domain: C-type lectin homology <LCH>  
 F:177-199/Region: PEST sequence  
 F:201-233/Region: PEST sequence  
 F:245-280/Domain: EGF homology <EG1>  
 F:288-323/Domain: EGF homology <EG2>  
 F:329-362/Domain: EGF homology <EG3>  
 F:369-404/Domain: EGF homology <EG4>  
 F:408-439/Domain: EGF homology <EG5>  
 F:445-480/Domain: EGF homology <EG6>  
 F:485-513/Region: PEST sequence  
 F:517-539/Domain: transmembrane #status predicted <TMN>  
 F:540-575/Domain: intracellular #status predicted <INT>  
 F:47,115,116,382,409/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:174,225,411,504/Binding site: carbohydrate (Thr) (covalent) #status predicted  
 F:245-256,252-265,267-280,288-296,292-308,310-322,329-340,336-349,351-362,369-378,374-38  
 F:334,498/Binding site: carbohydrate (Ser) (covalent) #status predicted  
 F:142/Modified site: erythro-beta-hydroxyasparagine (Asn) #status experimental  
 F:590,492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimen  
 Query Match 54.8%; Score 17; DB 1; Length 575;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 CXXXXXXX 12  
 DB 280 CTASATQSC 288  
 RESULT 24  
 T07953  
 lectin-like protein zsp2, zygote-specific - Chlamydomonas reinhardtii  
 C:Species: Chlamydomonas reinhardtii  
 C:Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 09-Jul-2004  
 C:Accession: T07953  
 R:Suzuki, L.; Woessner, J.P.; Uchida, H.; Kuriyama, H.; Takano, H.; Kuroiwa, H.; Yuasa,  
 submitted to the EMBL Data Library, March 1998  
 A:Description: A lectin-like protein mediates the assembly of the extracellular matrix c  
 A:Reference number: Z16243  
 A:Accession: T07953  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-582 <SUZ>  
 A:Cross-references: UNIPROT:O65153; UNIPARC:UPI00000454CF; EMBL:AF053099; NID:g2997677;  
 C:Genetics:  
 A:Gene: zsp2  
 A:introns: 27/1; 292/3  
 Query Match 54.8%; Score 17; DB 2; Length 582;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 CXXXXXXX 12  
 DB 257 CARSTTTC 265  
 RESULT 25  
 T07952

lectin-like protein, zygote-specific - Chlamydomonas reinhardtii  
 C:Species: Chlamydomonas reinhardtii  
 C:Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 09-Jul-2004  
 C:Accession: T07952  
 R:Stuiki, L.; Moesener, J.P.; Uchida, H.; Kuriyama, H.; Takano, H.; Kuroiwa, H.; Yuasa, submitted to the EMBL Data Library, March 1998  
 A:Description: A lectin-like protein mediates the assembly of the extracellular matrix  
 A:Reference number: Z16243  
 A:Accession: T07952  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-582 <SUZ>  
 A:Cross-references: UNIPROT:Q9SB11; UNIPARC:UPI000004454C; EMBL:AF053098; NID:g2997675;  
 C:Genetics:  
 A:Gene: zsp2  
 A:Introns: 27/1; 292/3

Query Match      54.8%; Score 17; DB 2; Length 582;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12  
 Db 257 CARSTTTC 265

RESULT 26  
 B25682  
 homeotic protein engrailed - fruit fly (Drosophila virilis)  
 C:Species: Drosophila virilis  
 C:Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 09-Jul-2004  
 C:Accession: B25682  
 R:Kasib, J.A.; Poole, S.J.; Wright, D.K.; O'Farrell, P.H.  
 EMBO J. 5, 3583-3589, 1986  
 A:Title: Sequence conservation in the protein coding and intron regions of the engrailed  
 A:Reference number: A91059; MUID:87161768; PMID:2881781  
 A:Accession: B25682  
 A:Molecule type: DNA  
 A:Residues: 1-584 <KAS>  
 A:Cross-references: UNIPROT:P09145; UNIPARC:UPI000012CA14; GB:X04727; NID:g9173; PIDN:CA  
 C:Genetics:  
 A:Gene: en  
 A:Cross-references: FlyBase:FBgn0013111  
 A:Introns: 470/1; 502/3  
 C:Superfamily: engrailed homeotic protein; homeobox homology  
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
 F:487-543/Domain: homeobox homology <HOX>

Query Match      54.8%; Score 17; DB 2; Length 584;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12  
 Db 384 GSASSSSGC 392

RESULT 27  
 S41011  
 hypothetical protein ZK757.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
 C:Accession: S41011  
 R:Thomas, K.  
 submitted to the EMBL Data Library, December 1993  
 A:Reference number: S41011  
 A:Accession: S41011  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-589 <THO>  
 A:Cross-references: UNIPROT:P34679; UNIPARC:UPI000013BC25; EMBL:Z29121; NID:g438366; PIDN:CA  
 C:Genetics:  
 A:Introns: 19/2; 54/3; 114/3; 146/3; 178/3; 209/1; 253/1; 302/3; 324/3; 352/1; 392/3; 46

C:Keywords: transmembrane protein

Query Match      54.8%; Score 17; DB 2; Length 589;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12  
 Db 100 CITSSSTTC 108

RESULT 28  
 A42086  
 CD30 antigen precursor - human  
 N:Alternate names: Ki-1 antigen; nerve growth factor receptor family member CD30  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: A42086  
 R:Durkop, H.; Latza, U.; Hummel, M.; Eitelbach, F.; Seed, B.; Stein, H.  
 Cell 68, 421-427, 1992  
 A:Title: Molecular cloning and expression of a new member of the nerve growth factor rec  
 A:Reference number: A42086; MUID:92154659; PMID:1310894  
 A:Accession: A42086  
 A:Molecule type: mRNA  
 A:Residues: 1-595 <DUR>  
 A:Cross-references: UNIPROT:P28908; UNIPARC:UPI0000000971; GB:M83554; NID:g180095; PIDN:CA  
 A:Experimental source: HUT-102 cell line  
 A:Note: sequence extracted from NCBI backbone (NCBIN:82088, NCBI:82090)  
 C:Genetics:  
 A:Gene: GDB:CD30; DLS166E  
 A:Cross-references: GDB:131547; OMIM:153243  
 A:Map position: 1p36-1p36  
 C:Keywords: glycoprotein; growth factor receptor; transmembrane protein  
 F:1-16/Domain: signal sequence #status predicted <Sig>  
 F:19-383/Domain: extracellular #status predicted <EXT>  
 F:384-407/Domain: transmembrane #status predicted <TM>  
 F:408-595/Domain: intracellular #status predicted <CYT>  
 F:101,276/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match      54.8%; Score 17; DB 2; Length 595;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12  
 Db 289 CATSATNSC 297

RESULT 29  
 D83286  
 hypothetical protein PA2886 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C:Accession: D83286  
 R:Stoyer, C.K.; Pham, X.O.; Bwlin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, J.N.; Folger, K.R.; Kae, A.; Lartig, K.; Lm,  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: D83286  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-600 <STO>  
 A:Cross-references: UNIPROT:Q9H2M1; UNIPARC:UPI00000C5861; GB:AE004714; GB:AE004091; NID:CA  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA2886

Query Match      54.8%; Score 17; DB 2; Length 600;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXXC 12  
DB 92 CAAALAAAC 100

## RESULT 30

T15408  
hypothetical protein C04F6.3 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C/Accession: T15408

R/Nhan, M.  
submitted to the EMBL Data Library, December 1995  
A/Description: The sequence of C. elegans cosmid C04F6.  
A/Reference number: Z18346

A/Accession: T15408  
A/Status: preliminary; translated from GB/EMBL/DDBL  
A/Molecule type: DNA  
A/Residues: 1-617 <NHA>  
A/Cross-references: UNIPROT:Q11174; UNIPARC:UPI0000060EB1; EMBL:U42835; NID:G125760; P1  
C/Genetics:  
A/Gene: CESP.C04F6.3  
A/Introns: 28/1; 66/2; 504/1

Query Match 54.8%; Score 17; DB 2; Length 617;  
Best Local Similarity 22.2%; Pred. No. 70;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXXC 12  
DB 524 CTTSTASGC 532

## RESULT 31

A36962  
laccase (EC 1.10.3.2) precursor - fungus (*Filobasidium floriforme*) (ATCC 34873)  
N/Alternate names: diphenol oxidase

C/Species: *Filobasidium neoforman*, *Cryptococcus neoformans*  
C/Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 31-Dec-2004  
C/Accession: A36962

R/Williamson, P.R.  
J. Bacteriol. 176, 656-664, 1994  
A/Title: Biochemical and molecular characterization of the diphenol oxidase of *Cryptococcus*  
A/Reference number: A36962; PMID:94131944; PMID:8300520  
A/Accession: A36962  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-624 <MIL>  
A/Cross-references: UNIPROT:Q9UQZ7; UNIPARC:UPI00001751D2; GB:I22866  
C/Genetics:  
A/Gene: CNLAC1  
A/Introns: 25/3; 69/3; 117/1; 137/3; 176/2; 204/2; 223/3; 238/3; 371/2; 391/2; 426/3; 55

C/Keywords: copper; glycoprotein; oxidoreductase

Query Match 54.8%; Score 17; DB 2; Length 624;  
Best Local Similarity 22.2%; Pred. No. 70;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXXC 12  
DB 243 CTATGSSSC 251

## RESULT 32

A41029  
integrin beta-8 chain precursor - human

C/Species: *Homo sapiens* (hmn)  
C/Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 31-Dec-2004  
C/Accession: A41029

R/Moyle, M.; Napier, M.A.; McLean, J.W.  
J. Biol. Chem. 266, 19650-19658, 1991  
A/Title: Cloning and expression of a divergent integrin subunit beta-8.  
A/Reference number: A41029; PMID:92011767; PMID:1918072

A/Accession: A41029  
A/Molecule type: mRNA  
A/Residues: 1-769 <MOY>

A/Cross-references: UNIPROT:P26012; UNIPARC:UPI000012DA14; GB:M73780; NID:G184520; PID:  
C/Superfamily: integrin, beta subunit; laminin-type BGF-like homology  
C/Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein

Query Match 54.8%; Score 17; DB 2; Length 769;  
Best Local Similarity 22.2%; Pred. No. 74;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXXC 12  
DB 47 CASSNASC 55

## RESULT 33

A46140  
diacylglycerol kinase (EC 2.7.1.107) - fruit fly (*Drosophila melanogaster*)  
C/Species: *Drosophila melanogaster*

C/Date: 22-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C/Accession: A46140

R/Masal, I.; Hosoya, T.; Kojima, S.; Hotta, Y.  
Proc. Natl. Acad. Sci. U.S.A. 89, 6030-6034, 1992

A/Title: Molecular cloning of a *Drosophila* diacylglycerol kinase gene that is expressed  
A/Reference number: A46140; PMID:92335231; PMID:1321433  
A/Accession: A46140

A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-791 <MAS>  
A/Cross-references: UNIPROT:Q01583; UNIPARC:UPI000016BB61; GB:D11120; NID:G217332; PID:  
A/Note: sequence extracted from NCBI backbone (NCBIN:108569, NCBI:P:108571)  
C/Genetics:  
A/Gene: FlyBase:Dgk  
A/Cross-references: FlyBase:FBgn0004568  
C/Superfamily: fruit fly diacylglycerol kinase; calmodulin repeat homology; protein kin

C/Keywords: BF hand; phosphotransferase

Query Match 54.8%; Score 17; DB 2; Length 791;  
Best Local Similarity 22.2%; Pred. No. 75;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXXC 12  
DB 18 CATPAAVAC 26

## RESULT 34

B46140  
diacylglycerol kinase (EC 2.7.1.107) - fruit fly (*Drosophila melanogaster*)

C/Species: *Drosophila melanogaster*  
C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 29-Sep-1999  
C/Accession: B46140

R/Masal, I.; Hosoya, T.; Kojima, S.; Hotta, Y.  
Proc. Natl. Acad. Sci. U.S.A. 89, 6030-6034, 1992

A/Title: Molecular cloning of a *Drosophila* diacylglycerol kinase gene that is expressed  
A/Reference number: A46140; PMID:92335231; PMID:1321433  
A/Accession: B46140

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-796 <MAS>  
A/Cross-references: UNIPARC:UPI00001756B6; GB:D11120  
C/Genetics:

A/Gene: FlyBase:Dgk  
A/Cross-references: FlyBase:FBgn0004568  
C/Superfamily: fruit fly diacylglycerol kinase; calmodulin repeat homology; protein kin  
C/Keywords: EF hand; phosphotransferase

Query Match 54.8%; Score 17; DB 2; Length 796;  
Best Local Similarity 22.2%; Pred. No. 75;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXXC 12

Db 18 CATAAAVAC 26

## RESULT 35

Integrin beta-5 chain precursor - human

C/Species: Homo sapiens (man)

C/Date: 14-Jun-1991 #sequence\_revision 14-Jun-1991 #text\_change 31-Dec-2004

C/Accession: A38308; A35775; S12534; S11708

R/McClean, J.W.; Vestral, D.J.; Cherech, D.A.; Bodary, S.C.

J Biol Chem. 265, 17126-17131, 1990

A/Title: cDNA sequence of the human integrin beta-5 subunit.

A/Reference number: A38308; M01D:91009141; PMID:2211615

A/Accession: A38308

A/Molecule type: mRNA

A/Residues: 1-799 <MCU>

A/Cross-references: UNIPROT:P18084; UNIPARC:UPI000040D5B; GB:J05633; NID:G186504; PIDN:

A/Note: parts of this sequence, including the amino end of the mature protein, were conf

R/Suzuki, S.; Huang, Z.S.; Tanihara, H.

Proc. Natl. Acad. Sci. U.S.A. 87, 5354-5358, 1990

A/Title: Cloning of an integrin beta subunit exhibiting high homology with integrin beta

A/Reference number: A35775; M01D:90319111; PMID:2371275

A/Accession: A35775

A/Molecule type: mRNA

A/Residues: 1-132, 'A', 194-644, 'L', 646-789, 793-799 <SUZ>

A/Cross-references: UNIPARC:UPI000016AADB; GB:M35011; NID:G184524; PIDN:AAAS2707.1; PID:

R/Ramaswamy, H.; Hemler, M.E.

EMBO J. 9, 1561-1568, 1990

A/Title: Cloning, primary structure and properties of a novel human integrin beta subunit

A/Reference number: S12534; M01D:90228356; PMID:2328726

A/Accession: S12534

A/Molecule type: mRNA

A/Residues: 1-644, 'L', 646-799 <RAM>

A/Cross-references: UNIPARC:UPI000012DA10; EMBL:X53002; NID:G33952; PIDN:CAA37188.1; PID:

C/Genetics:

A/Gene: GDB:ITG85

A/Cross-references: GDB:128005; OMIM:147561

A/Map position: 17q11-17qter

C/Suprafamily: Integrin, beta subunit; laminin-type EGF-like homology

C/Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-799/Product: integrin beta-5 chain #status experimental <MAT>

F:425-513/Domain: extracellular #status predicted <EXT>

F:463-513/Domain: laminin-type EGF-like homology <LEG>

F:720-744/Domain: transmembrane #status predicted <TM>

F:743-799/Domain: intracellular #status predicted <INT>

F:347,460,477,505,552,586,654,705/Binding site: carbohydrate (asn) (covalent) #status pr

Query Match 54.8%; Score 17; DB 2; Length 799;  
Best Local Similarity 22.2%; Pred. No. 75;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 28 CTSGSATSC 36

## RESULT 36

T05201

hypothetical protein F410.140 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C/Accession: T05201

R/Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Hobeisel, J.; Mew

submitted to the Protein Sequence Database, February 1999

A/Reference number: Z15402

A/Accession: T05201

A/Molecule type: DNA

A/Residues: 1-934 <BEV>

A/Cross-references: UNIPROT:Q9SMY8; UNIPARC:UPI000009D8D1; EMBL:AL035525

A/Experimental source: cultivar Columbia; BAC clone F410

A/Map position: 4  
A/Intons: 123/2; 203/3; 248/1; 281/2; 347/2; 437/3; 470/3; 500/2; 533/3; 590/3; 641/3;  
A/Note: F410.140

Query Match 54.8%; Score 17; DB 2; Length 934;  
Best Local Similarity 22.2%; Pred. No. 79;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 701 CTSATTASC 709

## RESULT 37

T41932

hypothetical protein U30 - human herpesvirus 7 (strain JI)

C/Species: human herpesvirus 7

A/Variety: strain JI

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C/Accession: T41932

R/Nicholas, J.

submitted to the EMBL Data Library, December 1995

A/Description: Determination and analysis of the complete nucleotide sequence of human h

A/Reference number: Z22022

A/Accession: T41932

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-938 <NIC>

A/Cross-references: UNIPROT:P52438; UNIPARC:UPI0000007D8; EMBL:U43400; PIDN:AAAC54692.1

A/Experimental source: strain JI

C/Genetics:

A/Note: U30

Query Match 54.8%; Score 17; DB 2; Length 938;  
Best Local Similarity 22.2%; Pred. No. 79;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 689 CSSTATTTC 697

## RESULT 38

T15881

hypothetical protein D1044.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C/Accession: T15881

R/Pauley, A.

submitted to the EMBL Data Library, June 1994

A/Description: The sequence of C. elegans cosmid D1044.

A/Reference number: Z18423

A/Accession: T15881

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1895 <PAU>

A/Cross-references: UNIPARC:UPI000004CAE8; EMBL:U00065; NID:G495681; PID:G495684; PIDN:

A/Experimental source: strain Bristol N2

C/Genetics:

A/Gene: CESP:D1044.3

A/Intons: 35/2; 63/2; 150/1; 191/3; 207/2; 405/3; 551/3; 629/1; 764/1; 933/1; 959/2; 10

Query Match 54.8%; Score 17; DB 2; Length 1895;  
Best Local Similarity 22.2%; Pred. No. 96;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 1233 CTSSSSSSC 1241

## RESULT 39

S50820



RESULT 44  
H66753  
propHage p12 protein 25 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: H66753  
R:Boletín, A.; Winkler, P.; Manger, S.; Jallón, O.; Malarne, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp  
A:Reference number: A86625; MWID:21235186; PMID:11337471  
A:Accession: H66753  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-87 <STO>  
A:Cross-references: UNIPROT:Q9CGR4; UNIPARC:UPI000006978; GB:AE005176; PID:g12723983; F  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: p1225

Query Match 51.6%; Score 16; DB 2; Length 87;  
Best Local Similarity 22.2%; Pred. No. 1.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 30 CTAKSLSC 38

RESULT 45  
J50036  
Clara cell 10K protein precursor - human  
N:Alternate names: urinary protein 1  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
C:Accession: J50036; P50309; A56890; I38397  
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Phillips, S.; Kennedy, A.L.; Anthony, J.; Squeglia  
Biochim. Biophys. Acta 950, 329-337, 1998  
A:Title: Amino-acid and cDNA nucleotide sequences of human Clara cell 10KDa protein.  
A:Reference number: J50036; MWID:89000784; PMID:3167058  
A:Accession: J50036  
A:Molecule type: mRNA  
A:Residues: 1-91 <STN>  
A:Cross-references: UNIPROT:P11684; UNIPARC:UPI000003690; GB:X13197; NID:g23131; PIDN:C  
A:Accession: P50309  
A:Molecule type: protein  
A:Residues: 22-23, 'X', 25-28, 'X', 30-31, 'X', 33-36 <SI2>  
A:Cross-references: UNIPARC:UPI0000177C39  
R:Bernard, A.; Roels, H.; Lauwerys, R.; Witters, R.; Gielens, C.; Soumilion, A.; Van De  
Clin. Chim. Acta 207, 239-249, 1992  
A:Title: Human urinary protein 1: evidence for identity with the Clara cell protein and  
A:Reference number: A56890; MWID:93009001; PMID:1395029  
A:Accession: A56890  
A:Molecule type: protein  
A:Residues: 22-45 <BER>  
A:Cross-references: UNIPARC:UPI0000036975  
A:Experimental source: urine  
A>Note: sequence extracted from NCBI backbone (NCBIP:119391)  
R:Hay, J.G.; Daniel, C.; Chu, C.; Crystal, R.G.  
Am. J. Physiol. 268, 565-575, 1995  
A:Title: Human CC10 gene expression in airway epithelium and subchromosomal locus sugges  
A:Reference number: I38397  
A:Accession: I38397  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-91 <RES>  
A:Cross-references: UNIPARC:UPI000003690; EMBL:U0110; NID:g45793; PIDN:AA8185.1; PI  
C:Comment: This protein consists of two identical polypeptides linked by two disulfide b  
C:Genetics:  
A:Gene: CC10  
A:Superfamily: uteroglobin  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-91/Product: Clara cell 10K protein #status experimental <MAT>

Query Match 51.6%; Score 16; DB 2; Length 91;  
Best Local Similarity 22.2%; Pred. No. 1.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 16 CSSASDTC 24

RESULT 46  
UGMS  
uteroglobin precursor - mouse  
N:Alternate names: CC10; Clara cell 10K protein precursor; Clara cell secretory protein  
C:Species: Mus musculus (house mouse)  
C:Date: 03-May-1994 #sequence\_revision 21-Jan-1997 #text\_change 09-Jul-2004  
C:Accession: A53025; A56656; I51925; S24783  
R:Striip, B.R.; Huffman, J.A.; Bohinski, R.J.  
Genomics 20, 27-35, 1994  
A:Title: Structure and regulation of the murine Clara cell secretory protein gene.  
A:Reference number: A53025; MWID:94292183; PMID:8020953  
A:Accession: A53025  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-96 <STR>  
A:Cross-references: UNIPROT:Q06318; UNIPARC:UPI00000011D0; GB:I24372; NID:g461147; PIDN:  
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Kennedy, A.L.  
Exp. Lung Res. 19, 67-75, 1993  
A:Title: Mouse Clara cell 10-KDa (CC10) protein: cDNA nucleotide sequence and molecular  
A:Reference number: A56656; MWID:93178380; PMID:8440203  
A:Accession: A56656  
A:Molecule type: protein  
A:Residues: 1-96 <STN>  
A:Cross-references: UNIPARC:UPI0000011D0; EMBL:X67702; NID:g49690; PIDN:CAA47936.1; PID  
A:Experimental source: lung  
A>Note: sequence extracted from NCBI backbone (NCBIP:126148)  
A:Note: parts of this sequence, including the amino end of the mature protein, were conf  
R:Margraf, L.R.; Finegold, M.J.; Stanley, L.A.; Major, A.; Watkins, H.K.; DeMayo, F.J.  
Am. J. Respir. Cell Mol. Biol. 9, 231-238, 1993  
A:Title: Cloning and tissue-specific expression of the cDNA for the mouse Clara cell 10K  
A:Reference number: I51925; MWID:94000840; PMID:8398159  
A:Accession: I51925  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-96 <RES>  
A:Cross-references: UNIPARC:UPI0000011D0; GB:L04503; NID:g20213; PIDN:AAA03625.1; PID  
C:Genetics:  
A:Insertions: 19/1; 81/3  
C:Complex: homodimer linked by two disulfide bonds  
C:Superfamily: uteroglobin  
C:Keywords: lung; steroid binding; uterus  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-96/Product: uteroglobin #status predicted <MAT>  
F:24/Disulfide bonds: interchain (to 90) #status predicted  
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 51.6%; Score 16; DB 1; Length 96;  
Best Local Similarity 22.2%; Pred. No. 1.2e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 16 CSSASDTC 24

RESULT 47  
A36581  
polychlorinated biphenyl-binding protein precursor - rat  
N:Alternate names: Clara cell 10K secretory protein  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 09-Jul-2004  
C:Accession: A36581; S10185; S21676  
R:Nordlund-Moeller, L.; Andersson, O.; Ahlgren, R.; Schilling, J.; Gyllner, M.; Gustafar

J. Biol. Chem. 265, 12690-12693, 1990  
A/Title: Cloning, structure, and expression of a rat binding protein for polychlorinated  
A/Reference number: A36581; PMID:90324266; PMID:2115524  
A/Accession: A36581  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-96 <NOR>  
A/Cross-references: UNIPROT:P17559; UNIPARC:UPI0000112E8B; GB:J05536; NID:G206039; PIDN:  
N18gen, G.; Wolf, M.; Karyal, S.L.; Singh, G.; Beato, M.; Suske, G.  
Nucleic Acids Res. 18, 2939-2946, 1990  
A/Title: Tissue-specific expression, hormonal regulation and 5'-flanking gene region of  
A/Reference number: S10185; PMID:90272398; PMID:2349092  
A/Accession: S10185  
A/Status: translation not shown  
A/Molecule type: DNA  
A/Residues: 1-18 <RHG>  
A/Cross-references: UNIPARC:UPI00001707CC; EMBL:X51318; NID:G55536; PIDN:CAA5701.1; PID  
R/Umland, T.C.; Swaminathan, S.; Purey, W.; Singh, G.; Fletcher, J.; Sax, M.  
J. Mol. Biol. 224, 441-448, 1992  
A/Title: Refined structure of rat Clara cell 17 kDa protein at 3.0 Å resolution.  
A/Reference number: S21676; PMID:92219263; PMID:1560460  
A/Contents: annotation, X-ray crystallography, 3.0 angstroms  
C/Superfamily: uteroglobin  
F/24/Disulfide bonds: interchain (to 24) #status experimental  
F/30/Disulfide bonds: interchain (to 24) #status experimental

Query Match 51.6%; Score 16; DB 2; Length 96;  
Best Local Similarity 22.2%; Pred. No. 1.2e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 16 CSASSSDIC 24

RESULT 48  
A10334  
conserved hypothetical protein YPO2745 [imported] - *Yersinia pestis* (strain CO92)  
C/Species: *Yersinia pestis*  
C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: A10334  
R/Parikh, J.; Wren, B.W.; Thomson, N.R.; Tildall, R.W.; Holden, M.T.G.; Prentice, M.B.;  
demo-Tariga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;  
J.L.M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A/Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A/Reference number: AB0001; PMID:21470413; PMID:11586360  
A/Accession: A10334  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-98 <KUR>  
A/Cross-references: UNIPROT:Q82D47; UNIPARC:UPI00000CD967; GB:AL590842; PIDN:CAC92984.1;  
C/Genetics:  
A/Gene: YPO2745  
C/Superfamily: C4-dicarboxylate carrier protein

Query Match 51.6%; Score 16; DB 2; Length 98;  
Best Local Similarity 22.2%; Pred. No. 1.2e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 8 CSAEETAC 16

RESULT 49  
S76421  
ferredoxin [2Fe-2S] - *Synechocystis* sp. (strain PCC 6803)  
C/Species: *Synechocystis* sp.  
A/Variety: PCC 6803  
C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C/Accession: S76421  
R/Kaneko, T.; Sato, S.; Kori, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasud,  
DNA Res. 3, 109-136, 1996  
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
S.  
A/Reference number: S74322; PMID:97061201; PMID:8905231  
A/Accession: S76421  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-103 <RAN>  
A/Cross-references: UNIPROT:P74449; UNIPARC:UPI00000D7139; EMBL:D90915; GB:AB013339; NIT  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C/Superfamily: ferredoxin [2Fe-2S]; ferredoxin [2Fe-2S] homology  
C/Keywords: 2Fe-2S; iron-sulfur protein; metalloprotein  
F/28-84/Domain: ferredoxin [2Fe-2S] homology <FER>  
F/43,48,51,83/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 51.6%; Score 16; DB 2; Length 103;  
Best Local Similarity 22.2%; Pred. No. 1.2e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 83 CAAYATSDC 91

RESULT 50  
J00863  
hypothetical 11.6k protein - *Escherichia coli* retron Ec67  
C/Species: *Escherichia coli* retron Ec67  
C/Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 09-Jul-2004  
C/Accession: J00863  
R/Hsu, M.Y.; Inouye, M.; Inouye, S.  
Proc. Natl. Acad. Sci. U.S.A. 87, 9454-9458, 1990  
A/Title: Retron for the 67-base multicopy single-stranded DNA from *Escherichia coli*: a  
A/Reference number: J00851; PMID:91067724; PMID:1701261  
A/Accession: J00863  
A/Molecule type: DNA  
A/Residues: 1-104 <RSU>  
A/Cross-references: UNIPROT:P21322; UNIPARC:UPI000013BF85; GB:M55249; NID:gl45143; PIDN  
A/Experimental source: *E. coli* strain Cl-1  
C/Genetics:  
A/Note: Insertion site is equivalent to 19 min of *E. coli* K12 genetic map  
C/Superfamily: *Escherichia coli* retron Ec67 hypothetical 11.6k protein

Query Match 51.6%; Score 16; DB 2; Length 104;  
Best Local Similarity 22.2%; Pred. No. 1.2e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 5 CSAEASHC 13

Search completed: January 4, 2006, 16:10:08  
Job time : 26.5217 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 97.9565 Seconds

(without alignments)  
108.037 Million cell updates/sec

Title: US-09-932-322-3

Perfect score: 31  
Sequence: 1 XXXXXXXXXXXXXXXX 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 1000 summaries

Database : uniprot\_05.80:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_trembl:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	58.1	144	2	061K6 DROME
2	18	58.1	155	2	P91214 CAEBL
3	18	58.1	269	2	07PR07 ANOQA
4	18	58.1	281	2	08GZ05 MAIZE
5	18	58.1	281	2	08LXK0 MAIZE
6	18	58.1	347	2	08G550 ORYSA
7	18	58.1	416	2	09ST87 ORYSA
8	18	58.1	468	2	07XST8 ORYSA
9	18	58.1	470	2	09LFB4 ARATH
10	18	58.1	504	2	09XZ88 LEIMA
11	18	58.1	510	2	08S4F6 ARATH
12	18	58.1	513	2	07XZ21 XENLA
13	18	58.1	518	2	0640H3 XENLA
14	18	58.1	569	2	07OKT3 GIALA
15	18	58.1	684	2	07XV79 ORYSA
16	18	58.1	772	1	DLIA_BRAAR
17	18	58.1	828	2	04WY29 ASPFU
18	18	58.1	1013	2	04SR96 TETNG
19	18	58.1	1031	2	04QRF5 LEIMA
20	18	58.1	1107	2	05CV70 CRYPV
21	18	58.1	1199	1	E7SBC DROME
22	18	58.1	1355	1	E7SBA DROME
23	18	58.1	1365	2	09VU79 DROME
24	18	58.1	1412	1	E7SBB DROME
25	18	58.1	1793	2	09W451 DROME
26	18	58.1	1836	2	08MP07 DROME
27	18	58.1	1894	2	04OS15 LEIMA
28	18	58.1	2087	2	04S488 TETNG
29	17	54.8	50	2	09L981 VIBCH
30	17	54.8	71	2	04RA29 TETNG
31	17	54.8	96	2	08VD96 MESAU

32	17	54.8	108	2	04PAH6 USTWA	04PAH6 ustlago ma
33	17	54.8	108	2	09ZSP6 ARATH	09ZSP6 arabidopsis
34	17	54.8	115	2	072D14 DESVH	072D14 desulfovibr
35	17	54.8	117	2	06VZRI CNPV	06VZRI canarypox v
36	17	54.8	119	2	04TWL2 ASF	04TWL2 african swi
37	17	54.8	119	2	04TWL3 ASF	04TWL3 african swi
38	17	54.8	119	2	04TWL5 ASF	04TWL5 african swi
39	17	54.8	119	2	04TWL6 ASF	04TWL6 african swi
40	17	54.8	119	2	04TWL7 ASF	04TWL7 african swi
41	17	54.8	119	2	04TWL8 ASF	04TWL8 african swi
42	17	54.8	119	2	04TWL9 ASF	04TWL9 african swi
43	17	54.8	119	2	04TMO ASF	04TMO african swi
44	17	54.8	125	2	05YPN5 NOCPA	05YPN5 nocardia fa
45	17	54.8	133	2	04TZU4 TETNG	04TZU4 tetraodon n
46	17	54.8	134	2	06ZRF5 HUMAN	06ZRF5 homo sapien
47	17	54.8	147	2	09CMG3 MOUSE	09CMG3 mus musculu
48	17	54.8	155	2	06P340 ORYSA	06P340 oryza sativ
49	17	54.8	155	2	04TWL4 ASF	04TWL4 african swi
50	17	54.8	157	1	CI025 HUMAN	081W50 homo sapien
51	17	54.8	157	1	CI025 MOUSE	09d772 mus musculu
52	17	54.8	157	2	04O6M1 LEIMA	04G6M1 leishmania
53	17	54.8	160	2	04J116 AZOVI	04J116 azobacter
54	17	54.8	163	2	09B1M0 GIALA	09B1M0 giardia lam
55	17	54.8	166	2	05TS94 HUMAN	05TS94 homo sapien
56	17	54.8	168	2	06K5J0 ORYSA	06K5J0 oryza sativ
57	17	54.8	170	2	08PRC3 XANAC	08PRC3 xanthomonas
58	17	54.8	171	2	04LIS2 YENTR	04LIS2 salmonella
59	17	54.8	171	2	093GN6 SALTU	093GN6 salmonella
60	17	54.8	172	2	05TS90 HUMAN	05TS90 homo sapien
61	17	54.8	172	2	06K2N4 ORYSA	06K2N4 oryza sativ
62	17	54.8	175	2	09VS07 DROME	09VS07 drosophila
63	17	54.8	183	2	05TS95 HUMAN	05TS95 homo sapien
64	17	54.8	186	1	CI025 BRAAR	06d670 brachydanio
65	17	54.8	191	2	05OZH7 BRAAR	05OZH7 brachydanio
66	17	54.8	199	2	07F588 CAEBL	07F588 caenorhabdi
67	17	54.8	202	2	06FX59 CANGA	06FX59 canidia gla
68	17	54.8	204	2	06YU14 ORYSA	06YU14 oryza sativ
69	17	54.8	207	2	08N1Y5 HUMAN	08N1Y5 homo sapien
70	17	54.8	212	1	YN87 YEAST	PS3719 saccharomyc
71	17	54.8	214	2	053846 MYCTU	053846 mycobacteri
72	17	54.8	214	2	07U158 MYCBO	07U158 mycobacteri
73	17	54.8	222	2	04VSY8 DROME	04VSY8 drosophila
74	17	54.8	225	2	09LJB7 ARATH	09LJB7 arabidopsis
75	17	54.8	228	2	04VSS4 DROME	04VSS4 drosophila
76	17	54.8	244	2	09Y9Q2 ABRPE	09Y9Q2 aceryprum p
77	17	54.8	247	2	06AP27 DEBHA	06AP27 debaryomyce
78	17	54.8	248	2	09USS9 SCRPO	09USS9 schizosacch
79	17	54.8	254	2	06WMS9 BRAAR	06WMS9 brachylosto
80	17	54.8	257	2	06ZDL1 ORYSA	06ZDL1 oryza sativ
81	17	54.8	259	2	083212 TRBPA	083212 treponema p
82	17	54.8	262	1	W0X3A MAIZE	07U0V1 zea mays (m
83	17	54.8	265	1	W0X3B MAIZE	06S313 zea mays (m
84	17	54.8	273	2	07SH84 ORYSA	07SH84 oryza sativ
85	17	54.8	281	2	057NZ1 SALCH	057NZ1 salmonella
86	17	54.8	281	2	05PHQ3 SALPA	05PHQ3 salmonella
87	17	54.8	281	2	08ZT96 SALTU	08ZT96 salmonella
88	17	54.8	281	2	08ZP73 SALTU	08ZP73 salmonella
89	17	54.8	283	2	07B973 MYCTU	07B973 mycobacteri
90	17	54.8	286	2	06GQD6 XENLA	06GQD6 xenopus lae
91	17	54.8	290	2	06A111 PROAC	06A111 propionibac
92	17	54.8	291	2	08W318 PROBI	08W318 vltis labru
93	17	54.8	293	1	XTX31 ARATH	093046 arabidopsis
94	17	54.8	298	2	08Z599 ARATH	08Z599 arabidopsis
95	17	54.8	305	2	04LMB9 GBRBK	04LMB9 burholderi
96	17	54.8	305	2	067L98 SYMTH	067L98 symbiodacte
97	17	54.8	306	2	07F8S0 ORYSA	07F8S0 oryza sativ
98	17	54.8	308	2	06ZJ81 BURMA	06ZJ81 burholderi
99	17	54.8	308	2	05GMI3 CAEBL	05GMI3 caenorhabdi
100	17	54.8	309	2	08GY34 ARATH	08GY34 arabidopsis
101	17	54.8	309	2	04NUR5 GDELT	04NUR5 anaeromyxob
102	17	54.8	310	2	05N9V7 ORYSA	05N9V7 oryza sativ
103	17	54.8	311	2	06EP90 ORYSA	06EP90 oryza sativ
104	17	54.8	311	2	063SX3 BURPS	063SX3 burholderi















```

981 16 51.6 571 2 Q6QH6_MOUSE Q6qih6_mus_musculu
982 16 51.6 571 2 Q5EG07_1CTPU Q5eg07_ictalunus p
983 16 51.6 572 2 Q9J8B6_9NCL Q9j8b6_spodoptera
984 16 51.6 574 2 Q828T3_STRAM Q828t3_streptomyc
985 16 51.6 575 2 Q7QNV8_GIALA Q7qnv8_giardia lam
986 16 51.6 575 2 Q7XNU7_ORYSA Q7xnu7_oryza sativ
987 16 51.6 577 2 Q7UX93_RHOBA Q7ux93_rhodoptrell
988 16 51.6 578 2 Q7SG68_NEUCR Q7sg68_neutrospora
989 16 51.6 581 2 Q6WMM9_BORCI Q6wmm9_botrytis ci
990 16 51.6 582 2 Q53PZ9_ORYSA Q53pz9_oryza sativ
991 16 51.6 587 1 SBU10_CAEEL Q93794_caenorhabdi
992 16 51.6 588 2 Q6GUE9_TOBAC Q6gue9_nicotiana t
993 16 51.6 588 2 Q5UX66_BACFN Q5ux66_bacteroides
994 16 51.6 588 2 Q64XS9_BACFR Q64xs9_bacteroides
995 16 51.6 592 2 Q60018_PICAN Q60018_pichia angu
996 16 51.6 592 2 Q98H21_RHILQ Q98h21_rhizobium l
997 16 51.6 595 2 Q9CZK3_NEUCR Q9czk3_neutrospora
998 16 51.6 595 2 Q5YXL7_HUMAN Q5yxl7_homo sapien
999 16 51.6 596 1 HMDH1_SOLTU P48020_solanum tub
1000 16 51.6 596 2 Q4K8Z7_PSEPS Q4k8z7_pseudomonas

```

## ALIGNMENTS

```

RESULT 1
06IKH6 DROME PRELIMINARY; PRT; 144 AA.
ID 06IKH6
AC 06IKH6
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HDCl2427.
GN ORFNames=HDC12427;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;
RA Hild M., Beckmann B., Haas S.A., Koch B., Solovjev V., Bueold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Hottel J.D.,
RA Pao R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome."
RL Genome Biol. 5:RSEARCH0003.1-RESEARCH0003.17(2003).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ child parry annotation (TPA) entry.
DR EMBL; BK002390; DAA03896.1; -; Genomic DNA.
SQ SEQUENCE 144 AA; 15277 MW; 1E4588BC183D5F2 CRC64;

```

```

Query Match 58.1%; Score 18; DB 2; Length 144;
Best Local Similarity 22.2%; Pred. No. 27;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 4 CXXXXXXXXC 12
DB 37 CATTASASC 45

RESULT 2
P91214 CAEEL PRELIMINARY; PRT; 155 AA.
ID P91214
AC P91214
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE C-type lectin protein 45.
GN Name=clec-45; ORFNames=F07C4.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;

```

```

OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=Bristol N2;
RC MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; U80023; AAG24036.1; -; Genomic DNA.
DR PIR; T28945; T28945.
DR Ensembl; F07C4.2; Caenorhabditis elegans.
DR Wormbase; WBGene00017199; F07C4.2.
DR WormPep; F07C4.2; CE09201.
DR GO; GO:0005529; P.sugar binding; IEA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; Lectin_C; 1.
DR SMART; SM00034; CLECT_1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW Complete proteome; Lectin.
SQ SEQUENCE 155 AA; 16285 MW; 956F0F4FDCB0F93 CRC64;

```

```

Query Match 58.1%; Score 18; DB 2; Length 155;
Best Local Similarity 22.2%; Pred. No. 28;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 4 CXXXXXXXXC 12
DB 92 CTTASSSSC 100

RESULT 3
Q7PRQ7 ANOGA PRELIMINARY; PRT; 469 AA.
ID Q7PRQ7
AC Q7PRQ7
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000001657 (Fragment).
GN ORFNames=ENSANGG00000001387;
OS Anopheles gambiae str. PE8T.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxId=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PE8T.
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PE8T.
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008647; BAA06779.3; -; Genomic DNA.
FT NON_TER 1
FT NON_TER 269
SQ SEQUENCE 269 AA; 24228 MW; 25BBPF71FD71F1F2 CRC64;

```

```

Query Match 58.1%; Score 18; DB 2; Length 269;
Best Local Similarity 22.2%; Pred. No. 34;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```



```

RESULT 4
08GZ05_MAIZE PRELIMINARY; PRT; 281 AA.
ID 08GZ05;
AC 08GZ05;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
GN DREB-like protein.
GN Name=drebl;
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxId=4577;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Qin F., Li J., Zhao J., Chen S.-Y., Liu Q.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF448789; AAN7673.1; -; mRNA.
DR HSSP; 080337; ZGCC.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2; 1.
DR PRINTS; PR00367; ETHRSPELEMT.
DR ProDom; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.
SQ SEQUENCE 281 AA; 29346 MW; A0902A7FFE0AB978 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 281;
Best Local Similarity 22.2%; Pred. No. 35;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12
Db 11 CSATTSSC 19

RESULT 5
08LKK0_MAIZE PRELIMINARY; PRT; 281 AA.
ID 08LKK0;
AC 08LKK0;
DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE DRE binding factor 2.
GN Name=dre2;
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxId=4577;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=22058745; PubMed=12061899;
RA Kizis D., Pages M.;
RT "Maize DRE-binding proteins DBF1 and DBF2 are involved in rab17
RT regulation through the drought-responsive element in an ABA-dependent
RT pathway."
RT Plant J. 30:679-689(2002).
RL EMBL; AF493799; AAM80485.1; -; mRNA.
DR HSSP; 080337; ZGCC.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2; 1.
DR PRINTS; PR00367; ETHRSPELEMT.
DR ProDom; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.

```

```

SQ SEQUENCE 281 AA; 29332 MW; CBE02A7FFE0AB979 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 281;
Best Local Similarity 22.2%; Pred. No. 35;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12
Db 11 CSATTSSC 19

RESULT 6
08Q550_9BETA PRELIMINARY; PRT; 347 AA.
ID 08Q550_9BETA PRELIMINARY; PRT; 347 AA.
AC 08Q550;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE UL38.
OS Fongine herpesvirus 4 (Chimpanzee cytomegalovirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxId=188763;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22421467; PubMed=12533697; DOI=10.1099/vir.0.18606-0;
RA Davison A.J., Dolan A., Akter P., Addison C., Dargan D.J.,
RA Alencor D.J., McGeoch D.J., Hayward G.S.;
RT "The human cytomegalovirus genome revisited: comparison with the
RT chimpanzee cytomegalovirus genome."
RL J. Gen. Virol. 84:17-28(2003).
DR EMBL; AF480884; AAM00688.1; -; Genomic DNA.
SQ SEQUENCE 347 AA; 39141 MW; 34B18107D501982 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 347;
Best Local Similarity 22.2%; Pred. No. 38;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12
Db 312 CSSTSSSTC 320

RESULT 7
09ST87_ORYSA PRELIMINARY; PRT; 416 AA.
ID 09ST87_ORYSA PRELIMINARY; PRT; 416 AA.
AC 09ST87;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DE 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE CAA103710.1 protein.
GN Name=q3037.10;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxId=4530;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA RC STRAIN=DNA;
RA Hong G., Chen Z.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ245900; CAB53483.1; -; Genomic DNA.
DR Gramene; 09ST87; -.
DR GO; GO:0004629; P:phospholipase C activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR000909; PI_PL_C_X.
DR PROSITE; PS50007; PIPLC_X DOMAIN; 1.
SQ SEQUENCE 416 AA; 45736 MW; 332A261949FFE6DD CRC64;

Query Match 58.1%; Score 18; DB 2; Length 416;
Best Local Similarity 22.2%; Pred. No. 41;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 4 CXXXXXXC 12  
DB 34 CTASASASC 42

RESULT 8  
Q7XST8\_ORYSA PRELIMINARY; PRT; 468 AA.

AC Q7XST8;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DE OSUNBA0039K24.18 protein.  
GN Name=OSUNBA0039K24.18;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]

RP NUCLEOTIDE SEQUENCE.  
RX PubMed:12447439; DOI=10.1038/nature01183;  
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,  
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,  
RA Weng Q., Zhang L., Lu Y., Mu Y., Lu Y., Zhang L.S., Yu Z., Fan D.,  
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,  
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yan H.,  
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia T., Zhang Y.,  
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,  
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,  
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,  
RA Han B.;  
RT "Sequence and analysis of rice chromosome 4.";  
RL Nature 420:316-320 (2002).  
DR EMBL; AL606637; CAB01799.2; -; Genomic\_DNA.  
DR Gramene; Q7XST8; -;  
DR GO; GO:0004629; P:phospholipase C activity; IEA.  
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
DR InterPro; IPR000909; PI PLC\_X.  
DR PROSITE; PS50007; PIP2C\_X DOMAIN; 1.  
SQ SEQUENCE 468 AA; 48671 MW; 49D661AA65338A33 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 468;  
Best Local Similarity 22.2%; Pred. No. 43;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 34 CTASASASC 42

RESULT 9  
Q9LFB4\_ARATH PRELIMINARY; PRT; 470 AA.

AC Q9LFB4;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DE Hypothetical protein F7J8\_200.  
GN Name=F7J8\_200;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC euroside II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;

RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Bevan M., Zimmermann W., Gruenisen A., Wambutt R., Kalicki J.,  
RA Wohldmann P., Smith A., Bancroft I., Mewes H.W., Rudd S., Lemcke K.,  
RA Mayer K.F.X.;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.  
RA EU Arabidopsis sequencing project;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL137189; CAB69850.1; -; Genomic\_DNA.  
DR PIR; T45962; T45962.  
DR GO; GO:0009058; P:photosynthesis; IEA.  
DR InterPro; IPR001296; Glyco\_transf\_1.  
DR Pfam; PF00534; Glycoe\_transf\_1; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 470 AA; 51797 MW; A7CB8BDAA6EF5BC6 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 470;  
Best Local Similarity 22.2%; Pred. No. 43;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 23 CASSATSC 31

RESULT 10  
Q9XZX8\_LEIMA PRELIMINARY; PRT; 504 AA.

AC Q9XZX8;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DE Possible surface antigen.  
GN Name=I302.01;  
OS Leishmania major.  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5664;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Friedlin;  
RX MEDLINE=98146435; PubMed=9477341;  
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
RA Smith D.F.;  
RT "A physical map of the Leishmania major Friedlin genome.";  
RL Genome Res. 8:135-145 (1998).  
RN [2]

RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=Friedlin;  
RA Ivens A.C., Murphy L., Quail M., Harris D., Oliver K., Lawson D.,  
RA Rajadream M.A., Barrell B.G.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL389894; CAC22669.1; -; Genomic\_DNA.  
DR InterPro; IPR006058; 2Fe2S\_f4\_BS.  
DR InterPro; IPR006212; Purin\_repeat.  
DR InterPro; IPR006210; IEGF\_repeat.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR SMART; SM00181; EGF; 3.  
DR SMART; SM00261; FU; 5.  
DR PROSITE; PS00197; 2FE2S\_FERREDOXIN; UNKNOWN 3.  
DR PROSITE; PS01248; LAMININ TYPE EGF; UNKNOWN 1.  
SQ SEQUENCE 504 AA; 52823 MW; B7545059F5959263 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 504;  
Best Local Similarity 22.2%; Pred. No. 44;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 161 CASTTASTC 169

RESULT 11

Q8S4F6\_ARATH PRELIMINARY; PRT; 510 AA.

AC Q8S4F6;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)

DE Sulfolipid synthase (Hypothetical protein At5g01220).  
 GN Name=SQD2; Synonyms=At5g01220;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RX PubMed=11960029; DOI=10.1073/pnas.082696499;  
 RP NUCLEOTIDE SEQUENCE.  
 RA Yu B., Xu C., Benning C.;  
 RT "Arabidopsis disrupted in SQD2 encoding sulfolipid synthase is  
 impaired in phosphate-limited growth."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:5732-5737(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Yamada K., Chan W.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,  
 RA Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,  
 RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,  
 RA Palm C.J., Shin P., Southwick A., Tripp M.G., Wu T., Davis R.W.,  
 RA Ecker J.R., Theologis A.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF454354; AAM18913.1; -, mRNA.  
 DR EMBL; BT005796; AA064198.1; -, mRNA.  
 DR GO; GO:0009058; P:biosynthesis; IEA.  
 DR InterPro; IPR001296; Glyco trans.1.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 DR Hypothetical protein.  
 KW SEQUENCE 510 AA; 56630 MW; 20E57D318DBED68 CRC64;  
 SQ

Query Match 58.1%; Score 18; DB 2; Length 510;  
 Best Local Similarity 22.2%; Pred. No. 44;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 12  
 DB 23 CSSSSATSC 31

RESULT 12  
 Q7ZX21\_XENLA PRELIMINARY; PRT; 513 AA.  
 AC Q7ZX21;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE MGCS3520 protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RP NUCLEOTIDE SEQUENCE.  
 RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ushed T.B., Toshlyuk S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Mair M.A.,

RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative."  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC046254; AAH46254.1; -, mRNA.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR003323; OTU.  
 DR Pfam; PF02338; OTU; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS0802; OTU; 1.  
 SQ SEQUENCE 513 AA; 55868 MW; A8734AD6E86FE6A CRC64;  
 SQ

Query Match 58.1%; Score 18; DB 2; Length 513;  
 Best Local Similarity 22.2%; Pred. No. 44;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 12  
 DB 376 CSSATTAAC 384

RESULT 13  
 Q640H3\_XENLA PRELIMINARY; PRT; 518 AA.  
 AC Q640H3;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE LOC494714 protein.  
 GN Name=LOC494714;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RX PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RP NUCLEOTIDE SEQUENCE.  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative."  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ushed T.B., Toshlyuk S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RN NUCLEOTIDE SEQUENCE.  
 RP TISSUE=Kidney;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC082654; AA82654.1; -; mRNA.  
 DR InterPro; IPR003323; OTU.  
 DR Pfam; PF02338; OTU; 1.  
 DR PROSITE; PS50802; OTU; 1.  
 SQ SEQUENCE 518 AA; 56209 MW; 2099AB07C6C93EC1 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 518;  
 Best Local Similarity 22.2%; Pred. No. 45;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
 DB 381 CSATTAAC 389

RESULT 14  
 ID Q7OX73\_GIALA PRELIMINARY; PRT; 569 AA.  
 AC Q7OX73;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE GLP 77 40692 38983.  
 OS Giardia lamblia ATCC 50803.  
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
 OX NCBI\_TaxID=184922;  
 [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA STRAIN=WB C6;  
 RX MEDLINE=21167845; PubMed=11104758; DOI=10.1074/jbc.M006589200;  
 RA Abel E.S., Davids B.U., Robles L.D., Loflin C.E., Gillin F.D.,  
 RA Chakrabarti R.;  
 RT "Possible roles of protein kinase A in cell motility and excystation  
 RT of the early diverging eukaryote Giardia lamblia.";  
 RL J. Biol. Chem. 276:10320-10329(2001).  
 RN [2]  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=WB C6;  
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,  
 RA Olsen G.J., Sogin M.L.;  
 RT "Draft sequence of the Giardia lamblia genome.";  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL; AACB0100056; EAA39878.1; -; Genomic DNA.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0005506; F:iron ion binding; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001450; 4Fe4S\_Fe\_S\_Bd.  
 DR InterPro; IPR005127; 4Fe4S\_VSP.  
 DR Pfam; PF03302; VSP; 1.  
 DR PRINTS; PR00353; 4FE4SFDXIN.  
 KM Electron transport; Transport.  
 SQ SEQUENCE 569 AA; 59086 MW; EB0DF4E2B8CC3F20 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 569;  
 Best Local Similarity 22.2%; Pred. No. 46;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
 DB 464 CSATTAAC 472

RESULT 15  
 ID Q7XV79\_ORYSA PRELIMINARY; PRT; 684 AA.  
 AC Q7XV79;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE OSUNB0076A22.6 protein.  
 GN Name=OSUNB0076A22.6;  
 GN Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 [1]  
 RN NUCLEOTIDE SEQUENCE.  
 RP MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;  
 RX Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,  
 RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Yang K., Yu S., Tang Y.,  
 RA Weng Q., Zhang L., Lu Y., Ma Y., Lu Y., Zhang L.S., Yu Z., Pan D.,  
 RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li Z., Hu H., Guan J.,  
 RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,  
 RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,  
 RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,  
 RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,  
 RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,  
 RA Han B.;  
 RT "Sequence and analysis of rice chromosome 4.";  
 RL Nature 420:316-320(2002).  
 RL EMBL; AL663016; CAD40795.3; -; Genomic DNA.  
 DR Gramene; Q7XV79; -;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000719; Prot. kinase.  
 DR InterPro; IPR008271; Ser\_Thr\_pkin.AS.  
 DR InterPro; IPR002290; Ser\_Thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot. kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TyrKc; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN; 1.  
 SQ SEQUENCE 684 AA; 75877 MW; 8CB1910F1409B14 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 684;  
 Best Local Similarity 22.2%; Pred. No. 50;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 16  
 ID DILA\_BRARE STANDARD; PRT; 772 AA.  
 AC GED18; 057462;  
 DT 13-SEP-2005 (Rel. 48, Created)  
 DT 13-SEP-2005 (Rel. 48, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Delta-like protein A precursor (Delta protein).  
 GN Name=dla;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12

Db 740 CASASTSAC 748

RESULT 17

Q4WY29\_ASPFU PRELIMINARY; PRT; 828 AA.

ID Q4WY29;

AC Q4WY29;

DT 13-SEP-2005 (TReMBLrel. 31, Created)

DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)

DE Repressor protein.

GN ORFNames=Afu3g14830;

OS Aspergillus fumigatus Af293.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Aspergillales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

NCBI\_TaxId=330879;

OK NCB1\_TaxId=330879;

RM NUCLEOTIDE SEQUENCE.

RP STRAIN=Af293;

RC Nierman W., Pain A., Anderson M.J., Mortman J., Kim H., Stanley J., Artzy A., Berriman M., Abe K., Archer D.B., Berneteo C., Bennett J., Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S., Farman M., Fedorova N., Fedorova N., Fielding T.V., Fischer R., Foster N., Frazer A., Garcia J.L., Garcia M.J., Gobie A., Goldman G.H., Goni K., Griffith-Jones S., Gwilliam R., Haas B., Haas H., Harris D., Horluchi H., Huang J., Humphrey S., Jimenez J., Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R., Kumagai T., Latton A., Latge J.-P., Li W., Lord A., Lu C., Majores W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M., Mouyia I., Mulligan S., Murphy L., O'Neill S., Paulsen I., Pevsner J., Petosa M., Price C., Pritchard B.L., Quail M.A., Rabinovich E., Rawlins N., Rajandream M.-A., Reichard U., Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M., Roming C.W., Rutter S., Salzberg S.L., Sanchez M., Sanchez-Ferreiro J.C., Saunders D., Seeger K., Squares R., Squares S., Tacheuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J., White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K., Machida M., Hall N., Barrell B., Denning D.W.;

RT "Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus.";

RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.

CC -! CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; AF01000002; FAL92104.1; -; Genomic DNA.

DR EMBL; AF01000002; 92421 MW; 2B1B831C5E4FF0C6 CRC64;

SQ SEQUENCE 828 AA; 92421 MW; 2B1B831C5E4FF0C6 CRC64;

Query Match 58.1%; Score 18; DB 2; Length 828;

Best Local Similarity 22.2%; Pred. No. 54;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12

Db 57 CTTTTTTC 65

RESULT 18

Q4SR96\_TETNG PRELIMINARY; PRT; 1013 AA.

ID Q4SR96;

AC Q4SR96;

DT 13-SEP-2005 (TReMBLrel. 31, Created)

DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)

DE Chromosome 11 SCAR14528, whole genome shotgun sequence.

DE (fragment).

GN ORFNames=GSTENG00014015001;

OS Tetradodon nigroviridis (Green puffer).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthopterygii; Perciformes; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

NCBI\_TaxId=99883;

OK NCB1\_TaxId=99883;

RM NUCLEOTIDE SEQUENCE.

RP Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N., Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A., Nicaud S., Jaffe D., Fisher S., Boute N., Castellano S., Dastiva C., Salanoubat M., Levy M., Boudet N., Vacherie B., Anthouard V., Jubin C., Castellic V., Katinka M., Vacherie B., Biemont C., Skalli Z., Catellico L., Poulin J., De Bernardis V., Craud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J., Parra G., Lardier G., Chappier C., McKernan K.J., McEwan P., Bosak S., Kellis M., Volff J.N., Gutig R., Zody M.C., Mesirov J., Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M., Landet V., Schachter V., Queller F., Sautin W., Scarpelli C., Winkler P., Lander E.S., Weissbach J., Roest Croillins H.;

RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype.";

RT Nature 431:946-957(2004).

RL Nature 431:946-957(2004).

RM NUCLEOTIDE SEQUENCE.

RP Genoscope; Whitehead Institute Centre for Genome Research; Submitted (Feb-2004) to the EMBL/GenBank/DBJ databases.

CC -! CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC EMBL; CA601014528; CAP96836.1; -; Genomic DNA.

DR InterPro; IPR000175; Na/nttran\_sympoort.

DR Pfam; PF00209; SNF; 2.

DR PRINTS; PR00176; NANESUMPOR.

DR ProDom; PD000448; Na/nttran\_sympoort; 2.

DR PROSITE; PS00610; NA\_NEUROTRAN\_SYM\_1; 1.

DR PROSITE; PS00754; NA\_NEUROTRAN\_SYM\_2; UNKNOWN\_1.

DR PROSITE; PS0267; NA\_NEUROTRAN\_SYM\_3; 2.

KT Symport; Transmembrane; Transport.

FT NON\_TER 1

FT NON\_TER 1

FT NON\_TER 1

SQ SEQUENCE 1013 AA; 110756 MW; 3D35CF87787652D CRC64;

Query Match 58.1%; Score 18; DB 2; Length 1013;

Best Local Similarity 22.2%; Pred. No. 58;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12

Db 728 CSASTSAC 736

RESULT 19

Q4QFB5\_LEIMA PRELIMINARY; PRT; 1031 AA.

ID Q4QFB5;

AC Q4QFB5;

DT 13-SEP-2005 (TReMBLrel. 31, Created)

DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=LmjF15.0700;

OS Leishmania major.

OC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Leishmania.

NCBI\_TaxId=5664;

OK NCB1\_TaxId=5664;

RM NUCLEOTIDE SEQUENCE.

RP STRAIN=Friedlin;

RC Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J., Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S., Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;

RL Submitted (Jun-2005) to the EMBL/GenBank/DBJ databases.

DR EMBL; CR005254; CAJ03294.1; -; Genomic DNA.

KW Hypothetical protein.

SQ SEQUENCE 1031 AA; 112673 MW; 65DB334CA5ECBEA CRC64;

Query Match 58.1%; Score 18; DB 2; Length 1031;

Best Local Similarity 22.2%; Pred. No. 59;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXC 12  
DB 791 CSSSSSSSAC 799

RESULT 20

Q5CV70\_CRYPV PRELIMINARY; PRT; 1107 AA.

AC Q5CV70; 10-MAY-2005 (TrEMBLrel. 30, Created)

DT 10-MAY-2005 (TrEMBLrel. 30, last sequence update)

DE Hypothetical protein.

GN ORFNames=cg48\_5210;

OS Crypsosporidium parvum.

OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

OC Crypsosporidiidae; Cryptosporidium.

OX NCBI\_TaxID=5807;

RN [1]

NUCLEOTIDE SEQUENCE.

RC STRAIN=Iowa type II;

RX PubMed=15044751; DOI=10.1126/science.1094786;

RA Abrahamson M.S., Templeton T.J., Enomoto S., Abraham J.E., Zhu G.,

LANCET C.A., Deng M., Liu C., Widmer G., Tzipori S., Buck G.A., Xu P.,

RA Banker A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,

RT "Complete genome sequence of the apicomplexan, *Cryptosporidium*

parvum.";

RL Science 304:441-445 (2004).

CC -1- CAUTION: The sequence shown here is derived from an

EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

preliminary data.

CC EMBL: AABE0100003; EAK83643.1; -; Genomic\_DNA.

DR InterPro: IPR012283; 6PGD C core.

DR InterPro: IPR011989; ARM-like.

DR InterPro: IPR012293; Globin-related.

DR InterPro: IPR001524; Glyco\_hydro\_6.

DR InterPro: IPR012287; Homodomain-rel.

DR InterPro: IPR001093; IMPH/GMPase.

DR InterPro: IPR008976; PLAT IM2.

DR InterPro: IPR000783; RNAPOL\_RPB5.

DR InterPro: IPR009032; Ypu\_Cyt.

DR InterPro: IPR011991; Wing\_hix\_DNA\_bd.

DR Pfam: PF00478; IMPDH\_1.

KW Hypothetical protein.

SQ SEQUENCE 1107 AA; 127343 MW; 4BAD9A836ECB2AF CRC64;

Query Match 58.1%; Score 18; DB 2; Length 1107;

Best Local Similarity 22.2%; Pred. No. 60;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXC 12

DB 185 CSSSSSSSAC 193

RESULT 21

E75BC DROME STANDARD; PRT; 1199 AA.

AC P17671; 081081; 081083;

DT 01-AUG-1990 (Rel. 15, Created)

DT 13-SEP-2005 (Rel. 48, last sequence update)

DT 13-SEP-2005 (Rel. 48, last annotation update)

DE Ecdysone-induced protein 75B isoforms C/D (E75-A).

GN Name=Ecdysone-induced protein 75B isoforms C/D (E75-A).

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]  
NUCLEOTIDE SEQUENCE [mRNA], AND ALTERNATIVE SPLICING.

RC STRAIN=Cancon-S;

RX MEDLINE=90249727; PubMed=2110921;

RA Segaves W.A., Hognes D.S.;

RT "The E75 ecdysone-inducible gene responsible for the 75B early puff in

RT Drosophila encodes two new members of the steroid receptor

RT superfamily.";

RL Genes Dev. 4:204-219 (1990).

RN [2]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;

RA Adams M.D., Ceiniker S.B., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Abpayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck Y., Brokstein P., Brotier P.,

RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Casley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabrielian A.B., Gary N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Jarmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lascko P., Lei Y., Levitsky A.A., Li H., Li Z., Liang Y., Lin X.,

RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Put V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*.";

RL Science 287:2185-2195 (2000).

RN [3]

GENOME REANNOTATION, AND ALTERNATIVE SPLICING.

RX MEDLINE=22426069; PubMed=12537572;

RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.B.,

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,

RA Betencourt B.R., Ceiniker S.E., de Grey A.D.N.J., Drysdale R.A.,

RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,

RA Lewis S.E.;

RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a

RT systematic review.";

RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).

RN [4]

FUNCTION, DEVELOPMENTAL STAGE, AND INDUCTION.

RX MEDLINE=94038699; PubMed=8223281;

RA Hueb F., Ruiz C., Richards G.;

RT "Puffs and PCR: the in vivo dynamics of early gene expression during

RT ecdysone responses in *Drosophila*.";

RL Development 118:613-627 (1993).

CC -1- FUNCTION: Implicated in the regulation of ecdysone-triggered gene

hierarchies. Probably plays a key role in mediating the regulation



of the larval molt by 20-OH-ecdysone.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=4;  
 CC Name=C; Synonyms=E75A;  
 CC IsoId=P1671-1; Sequence=Displayed;  
 CC Name=A; Synonyms=E75B;  
 CC IsoId=P1672-1; Sequence=External;  
 CC Name=B; Synonyms=E75C;  
 CC IsoId=P13055-1; Sequence=External;  
 CC Name=D;  
 CC IsoId=P1671-2; Sequence=VSP 014915, VSP 014916;  
 CC Note=No experimental confirmation available;  
 CC -1- DEVELOPMENTAL STAGE: In mid instar larvae salivary glands, low  
 CC basal levels are observed in puff stage 1. Levels increase in late  
 CC larvae from puff stages 3-10, then decrease abruptly at stage 11.  
 CC In prepupae, isoform C is the predominant form during the  
 CC transition between puff stages 18-19. At puff stage 1, expression  
 CC is also present in the gut. By stage 3 it is present in the wing  
 CC disks, Malpighian tubules and the fat body. At stage 11,  
 CC expression is only present in the gut and wing disks.  
 CC -1- INDUCTION: The expression of this protein is developmentally  
 CC regulated and is correlated with the 20-OH-ecdysone induced  
 CC activity of puff 75B.  
 CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NRI  
 CC subfamily.  
 CC -1- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.  
 CC -1- CAUTION: Ref. 1 sequence differs from that shown due to a  
 CC frameshift in position 1164.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL, X51548; CAA35923.1; ALT FRAME: mRNA.  
 CC EMBL, AE003522; AAN11687.1; -; Genomic DNA.  
 CC EMBL, AE003522; AAN11689.1; -; Genomic DNA.  
 CC PIR, A34598; A34598.  
 CC HSSP, P20393; 1GA5.  
 CC SMR, P1671; 240-323.  
 CC TRANSFAC; T01367; -;  
 CC PLAYBASE; FBgn000568; E1p75B.  
 CC GO; GO:0004879; P11gand-dependent nuclear receptor activity; NAS.  
 CC GO; GO:0018990; Piecysis (sensu Insecta); IMP.  
 CC GO; GO:0035072; Piecysone-mediated induction of salivary gla. .; NAS.  
 CC GO; GO:0007553; P-regulation of ecdysteroid metabolism; IMP.  
 CC InterPro; IPR001628; Hrmn\_rcpt\_DNA\_bd.  
 CC InterPro; IPR000536; Hrmn\_rcpt\_lig\_bd.  
 CC InterPro; IPR001723; Stdhlm\_receptor.  
 CC InterPro; IPR000324; Vldt\_receptor.  
 CC Pfam; PF00105; Hormone\_recep; 1.  
 CC Pfam; PF00105; zf-C4; 1.  
 CC PRINTS; PR000398; STRODHORMNER.  
 CC PRINTS; PR00047; STRODHORMNER.  
 CC PRINTS; PR00350; VITAMINDR.  
 CC ProDom; PD000035; Znf\_C4steroid; 1.  
 CC SMART; SM00399; Znf\_C4; 1.  
 CC SMART; SM00430; HOL1; 1.  
 CC PROSITE; PS00031; NUCLEAR\_REC\_DBD\_1; 1.  
 CC PROSITE; PS01030; NUCLEAR\_REC\_DBD\_2; 1.  
 CC Alternative splicing; Developmental protein; DNA-binding;  
 CC Metal-binding; Nuclear protein; Receptor; Transcription;  
 CC Transcription regulation; Zinc; Zinc-finger.  
 CC KW DNA BIND 242 318  
 CC ZN\_FING 245 265 NR C4-type.  
 CC ZN\_FING 282 306 NR C4-type.  
 CC REGION 380 390 Ligand-binding (Potential).  
 CC COMPIAS 139 195 Gln/His-rich.  
 CC COMPIAS 721 756 Gln-rich.  
 CC COMPIAS 900 1103 Ser-rich.  
 CC VARSLIC 1 293 Missing (in isoform D).

FT VARSPLIC 294 315 /FTId=VSP 014915.  
 FT NNNRCYCKKCKICVAGMSRDA -> MGEELPILKILKGN  
 FT VYFNAP (in isoform D).  
 FT /FTId=VSP 014916.  
 SQ SEQUENCE 1199 AA; 128534 MW; E29372F5E295F595 CRC64;  
 Query Match 58.1%; Score 18; DB 1; Length 1199;  
 Best Local Similarity 22.2%; Pred. No. 63;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 CXXXXXXXC 12  
 Db 1070 CXXXXXXXC 1078  
 RESULT 22  
 ID E75BA DROME STANDARD; PRT; 1355 AA.  
 AC P1672; 081052;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 13-SEP-2005 (Rel. 48, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Ecdysone-induced protein 75B isoform A (E75-B).  
 GN Name=E1p75B; Synonyms=NR1D3; ORFNames=C8127;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxId=7227;  
 [1]  
 RN NUCLEOTIDE SEQUENCE [MRNA], AND ALTERNATIVE SPLICING.  
 RP STRAIN=canon-S;  
 RC MEDLINE=90249727, PubMed=2110921;  
 RX MEDLINE=90249727, PubMed=2110921;  
 RT "The E75 ecdysone-inducible gene responsible for the 75B early puff in  
 RT Drosophila encodes two new members of the steroid receptor  
 RT superfamily.";  
 RL Gene Dev. 4:204-219(1990).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celinker S.E., Holt R.A., Iwans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champo M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt D., Nelson C.R., Milos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mayhew A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doud L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.B., Garg N.S., Gelpi W.M., Glasser K.,  
 RA Glodak A., Gong F., Gorell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Hwang M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasok P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,



RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RA "The genome sequence of *Drosophila melanogaster*,"  
 RL Science 287:2185-2195(2000).  
 [3]  
 RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.  
 RA MEDLINE:22426069; PubMed:12537572.  
 RA Mista S., Crosby M.A., Mungall C.U., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kinkner J.S., Milburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Betencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter S., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.,  
 RA "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review,"  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 [4]  
 RP FUNCTION, DEVELOPMENTAL STAGE, AND INDUCTION.  
 RA MEDLINE:94038699; PubMed:8223281;  
 RA Hueb F., Ruiz C., Richards G.,  
 RT "Puffs and PCR: the in vivo dynamics of early gene expression during  
 RT ecdysone responses in *Drosophila*,"  
 RL Development 118:613-627(1993).  
 CC -1- FUNCTION: Implicated in the regulation of ecdysone-triggered gene  
 CC hierarchies. Probably plays a key role in mediating the regulation  
 CC of the larval molt by 20-OH-ecdysone.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=4;  
 CC Name=A; Synonyms=E75B;  
 CC IsoId=PI1672-1; Sequence=Displayed;  
 CC Name=C; Synonyms=E75A;  
 CC IsoId=PI1671-1; Sequence=External;  
 CC Name=B; Synonyms=E75C;  
 CC IsoId=PI1055-2; Sequence=External;  
 CC Name=D;  
 CC IsoId=PI1671-2; Sequence=External;  
 CC Note=No experimental confirmation available;  
 CC -1- DEVELOPMENTAL STAGE: In mid instar larvae salivary glands, levels  
 CC are low during puff stage 1, increase during puff stages 2-4 and  
 CC diminish from stage 5 onwards. In prepupae, isoform A is the  
 CC predominant form during puff stage 19 and the transition to stage  
 CC 20. By stage 3 it is present in the gut, Malpighian tubules and  
 CC the fat body, levels persist beyond stage 11.  
 CC -1- INDUCTION: The expression of this protein is developmentally  
 CC regulated and is correlated with the 20-OH-ecdysone induced  
 CC activity of puff 75B.  
 CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NRI  
 CC subfamily.  
 CC -1- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.  
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a  
 CC frameshift in position 1320.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL: X51549; CAA35924.1; ALT\_FRAME; mRNA.  
 CC EMBL: AE003522; AAN11688.1; -; Genomic\_DNA.  
 CC PIR: B34598; B34598.  
 CC HSSP: P20393; 1A6Y.  
 CC SMR: P1672; 415-480.  
 CC TRANSFAC: T01368; -;  
 CC Flybase: FBgn0000568; E1P75B.  
 CC GO: GO:0004879; F:ligand-dependent nuclear receptor activity; NMS.  
 CC GO: GO:0019990; P:ecdysone (sensa insecta); IMP.  
 CC GO: GO:0035072; P:ecdysone-mediated induction of salivary gla. .; NMS.

DR GO: GO:0007553; P:regulation of ecdysteroid metabolism; IMP.  
 DR InterPro: IPR001628; Hrm rcpt\_Lig bd.  
 DR InterPro: IPR000536; Hrm rcpt\_Lig bd.  
 DR InterPro: IPR000354; Involucrin\_rpt.  
 DR InterPro: IPR003079; KOR\_receptor.  
 DR InterPro: IPR001723; StdImm\_receptor.  
 DR Pfam: PF00104; Hormone\_recep; 1.  
 DR Pfam: PF00904; Involucrin; 1.  
 DR Pfam: PF00105; zf-C4; 1.  
 DR PRINTS: PR01293; KORNUCCEPTR.  
 DR PRINTS: PR00398; STRDHOMNER.  
 DR PRINTS: PR00047; STROIDPINGER.  
 DR ProDom: PD000035; ZnfC4Steroid; 1.  
 DR SMART: SM00430; HOL1; 1.  
 DR SMART: SM00399; ZNF C4; 1.  
 DR PROSITE: PS51030; NUCLEAR\_REC\_DBD\_2; 1.  
 KW Alternative splicing; Developmental protein; DNA-binding;  
 KW Metal-binding; Nuclear protein; Receptor; Transcription;  
 KW Transcription regulation; Zinc; Zinc-finger.  
 FT DNA\_BIND 384 474  
 FT ZN\_FING 387 421  
 FT COMPBIAS 438 457  
 FT COMPBIAS 6 324  
 FT COMPBIAS 877 912  
 FT COMPBIAS 1056 1259  
 FT CONFLICT 206 206  
 FT CONFLICT 244 244  
 FT CONFLICT 244 244  
 FT SEQUENCE 1355 AA; 147171 MW; 6E19BACEC56255 CRC64;  
 Query Match 58.1%; Score 18; DB 1; Length 1355;  
 Best Local Similarity 22.2%; Pred. No. 65;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 CXXXXXXC 12  
 DB 1226 CXXXXXXC 1234  
 RESULT 23  
 ID Q9VU79 DROME PRELIMINARY; PRT; 1365 AA.  
 AC Q9VU79;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE CG10231-PA.  
 GN Name=Pde11; ORFNames=CG10231;  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; *Drosophila*.  
 OC NCBI\_TaxId=7227;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE:20196006; PubMed:10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans G.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
 RA Suton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Foster K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,



RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Strupki M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskae R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
Wang Z.-Y., Massarman D.A., Weinstein G.M., Weisenbach J.,  
RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RA "The genome sequence of *Drosophila melanogaster*,"  
Science 287:2185-2195(2000).  
RL Science 287:2185-2195(2000).  
RN [3]  
RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
RA Beutenkourt B.R., Celniker S.E., de Grey A.D.N.U., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review,"  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [4]  
RP FUNCTION, DEVELOPMENTAL STAGE, AND INDUCTION.  
RX MEDLINE=94038699; PubMed=8223281;  
RA Huet F., Ruiz C., Richards G.;  
RT "Puffs and PCR: the in vivo dynamics of early gene expression during  
RT ecdysone responses in *Drosophila*,"  
RL Development 118:613-627(1993).  
CC -!- FUNCTION: Implicated in the regulation of ecdysone-triggered gene  
CC hierarchies. Probably plays a key role in mediating the regulation  
CC of the larval molt by 20-OH-ecdysone.  
CC -!- SUBCELLULAR LOCATION: Nuclear (potential).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=B; Synonyms=E75C;  
CC IsoId=PI3055-2; Sequence=Displayed;  
CC Name=C; Synonyms=E75A;  
CC IsoId=PI671-1; Sequence=External;  
CC Name=A; Synonyms=E75B;  
CC IsoId=PI672-1; Sequence=External;  
CC Name=D;  
CC IsoId=PI671-2; Sequence=External;  
CC Note=No experimental confirmation available;  
CC -!- DEVELOPMENTAL STAGE: In mid instar larvae salivary glands, levels  
CC increase during puff stage 1, then remain relatively constant  
CC until the premetamorphic pulse of ecdysone at puff stage 5. Levels  
CC increase again in late larvae at puff stages 9-10. At puff stage 1  
CC expression is also seen in the gut. Levels are low in the gut,  
CC Malpighian tubules, fat body and wing disks between stages 1 and  
CC 11.  
CC -!- INDUCTION: The expression of this protein is developmentally  
CC regulated and is correlated with the 20-OH-ecdysone induced  
CC activity of puff 75B.  
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NRI  
CC subfamily.  
CC -!- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.  
CC -!- CAUTION: Ref. 1 sequence differs from that shown due to frameshifts  
CC in positions 15, 543, 576, 593, 614 and 305.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC EMBL: X15586; CAA3361.1; ALT\_FRAME, mRNA.  
CC EMBL: AE003522; AAF49282.3; -f Genomic\_DNA.  
CC FIR: S05979; S05979.  
CC HSSP: P20393; 1GA5.  
CC SMR: P13055; 449-532.

DR Ensembl; CG8127; *Drosophila melanogaster*.  
DR FlyBase; FBgn000568; Eip75B.  
DR GO; GO:0004879; F. ligand-dependent nuclear receptor activity; NAS.  
DR GO; GO:0018990; P. ecdysis (sensu Insecta); IMP.  
DR GO; GO:0035072; P. regulation-mediated induction of salivary gla. . .; NAS.  
DR GO; GO:0005753; P. regulation of ecdysteroid metabolism; IMP.  
DR InterPro; IPR001628; Hrm rcpt DNA bd.  
DR InterPro; IPR000536; Hrm rcpt lig bd.  
DR InterPro; IPR001723; Strdim receptor.  
DR InterPro; IPR000324; Vld receptor.  
DR Pfam; PF00104; Hormone\_recep; 1.  
DR Pfam; PF00105; zt-C4; 1.  
DR PRINTS; PR00398; STRDHORMONER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR PRINTS; PR00350; VITAMINDR.  
DR ProDom; PD000035; Znf C4etereid; 1.  
DR SMART; SM00430; HOL1; 1.  
DR SMART; SM00399; Znf C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_REC DBD 1; 1.  
DR PROSITE; PS10030; NUCLEAR\_REC DBD 2; 1.  
DR KX Alternative splicing; developmental protein; DNA-binding;  
DR KW Metal-binding; Nuclear protein; Receptor; Transcription;  
DR KM Transcription regulation; Zinc; Zinc-finger.  
DR KW Transcription regulation; Zinc; Zinc-finger.  
DR FT DNA BIND 455 531 Nuclear receptor.  
DR FT ZN FING 458 478 NR C4-type 1.  
DR FT ZN FING 495 514 NR C4-type 2.  
DR FT REGION 589 797 ligand-binding (potential).  
DR FT COMEBIAS 6 10 poly-Ala.  
DR FT COMEBIAS 180 395 Gln-rich.  
DR FT COMEBIAS 265 298 Pro-rich.  
DR FT COMEBIAS 351 445 Ser-rich.  
DR FT COMEBIAS 934 969 Gln-rich.  
DR FT COMEBIAS 1113 1316 Ser-rich.  
DR FT CONFLICT 20 20 V -> F (1n Ref. 1).  
DR FT CONFLICT 183 183 Q -> QQ (1n Ref. 1).  
DR FT CONFLICT 337 342 Missing (1n Ref. 1).  
DR FT CONFLICT 1142 1142 S -> V (1n Ref. 1).  
DR FT CONFLICT 1216 1216 V -> L (1n Ref. 1).  
DR FT CONFLICT 1231 1231 Missing (1n Ref. 1).  
DR FT CONFLICT 1302 1302 N -> I (1n Ref. 1).  
DR SQ SEQUENCE 1412 AA; 151292 MW; 6F4C21B075443F8C CRC64;  
  
Query Match 58.1%; Score 18; DB 1; Length 1412;  
Best local similarity 22.2%; Pred. No. 66;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 4 CXXXXXXC 12  
DB 1283 CXXXXXXC 1291  
  
RESULT 25  
ID Q9W451 DROME PRELIMINARY; PRT; 1793 AA.  
AC Q9W451;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE CG4790-PA.  
GN Name=fs(1)M3; ORFNames=CG4790;  
OS *Drosophila melanogaster* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; *Drosophila*.  
OX NCBI\_TaxId=7227;  
ON [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miliot G.L.G.,  
 RA Abitl J.F., Abdayani A., An H.-J., Andrews-Framkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Bealey E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Bortova K.C., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng I., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Flischmann W.,  
 RA Foster C., Gabriellan A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasako P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclebo J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spler E., Spredling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195 (2000).

[2]

RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Murthy D.M., Nelson C.R.,  
 RA Paclebo J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers B.W., Gibbs R.A., Rubin G.M.,  
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
 RT *melanogaster* euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

[3]

RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.,  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RT a genomic perspective.";  
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
 [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter U., Russo S., Schroeder A.V., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.,  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RX Berkeley *Drosophila* Genome Project;  
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
 RA Hoskins R., Stapleton M., Paclebo J., Park S., Svirskas R., Smith E.,  
 RA Yu C., Rubin G.,

RT "Drosophila melanogaster release 4 sequence.";  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Flybase;  
 RL Submitted (MAR-2005) to the EMBL/Genbank/DBJ databases.  
 DR EMBL, AE003435; AAFA6107.1; -, Genomic DNA.  
 DR Ensembl, CG4790; Drosophila melanogaster.  
 DR Flybase, FBgn0005390; CG4790.  
 DR Flybase, FBgn0005390; fs(1)M3.  
 SQ SEQUENCE 1793 AA; 202233 MW; DDC38A764E301CD CRC64;  
 Query Match 58.1%; Score 18; DB 2; Length 1793;  
 Best Local Similarity 22.2%; Pred. No. 73;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 CXXXXXXKC 12  
 DB 291 CTTSSASASC 299  
 RESULT 26  
 OBP07 DROME PRELIMINARY; PRT; 1836 AA.  
 ID OBP07 DROME PRELIMINARY;  
 AC OBP07  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Poliole precursor.  
 GN Name=fs(1)M3; Synonyms=fs(1)ph;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=2195667; PubMed=11959640; DOI=10.1101/gad.223902;  
 RA Jimenez G., Gonzalez-Reyes A., Casanova U.,  
 RT "Cell surface proteolins Nasrat and Poliole stabilize the Torso-like  
 RT extracellular determinant in *Drosophila* oogenesis.";  
 RL Genes Dev. 16:913-918(2002).  
 DR Ensembl, AJ487522; CAD31790.1; -, mRNA.  
 DR Ensembl, CG4790; Drosophila melanogaster.  
 DR Flybase, FBgn0005390; CG4790.  
 DR Flybase, FBgn0005390; fs(1)M3.  
 FT SIGNAL.  
 KW Signal.  
 FT CHAIN 1 25 Potential.  
 FT CHAIN 26 1836 Poliole.  
 SQ SEQUENCE 1836 AA; 207449 MW; 4B0FC64DCDBF52F CRC64;  
 Query Match 58.1%; Score 18; DB 2; Length 1836;  
 Best Local Similarity 22.2%; Pred. No. 74;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 CXXXXXXKC 12  
 DB 291 CTTSSASASC 299  
 RESULT 27  
 Q40515 LEIMA PRELIMINARY; PRT; 1894 AA.  
 ID Q40515 LEIMA PRELIMINARY;  
 AC Q40515;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Calpain-like cysteine peptidase, putative (Cysteine peptidase, clan  
 DE ca, family c2, putative).  
 GN ORFNames=Lmjf32.0970;  
 GN Leishmania majOr.  
 OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 NCBI\_TaxID=5664;

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Bertman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrett B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005269; CAJ08617.1; -; Genomic DNA.
SQ SEQUENCE 1894 AA; 199683 MW; 3F71DE2F53BP95BB CRC64;

Query Match
Best Local Similarity 22.2%; Score 18; DB 2; Length 1894;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 1785 CAAAAAAC 1793

RESULT 28
Q4S488.TETNG PRELIMINARY; PRT; 2087 AA.
AC Q4S488;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAP14743, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00024299001;
OC Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Tetraodontiformes;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxId=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Coataz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bernardis V.,
RA Crnaud C., Duprat S., Broctier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Winkler P., Schachter V., Queller F., Saurin W., Scapellato C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG GenomeScope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; CAAB01014743; CAG04544.1; -; Genomic DNA.
DR EMBL; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_CA_bd.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR012680; Laminin_G_2.
DR Pfam; PF00028; Cadherin_8.
DR Pfam; PF01049; Cadherin_C_1.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF02210; Laminin_G_2; 2.

```

```

DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 9.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00282; LamG; 2.
DR PROSITE; PS00232; CADHERIN_1; 6.
DR PROSITE; PS50268; CADHERIN_2; 9.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS00071; GAPDH; UNKNOWN_1.
DR PROSITE; PS50025; LM_G_DOMAIN; 2.
KW Calcium; EGF-like domain; Transmembrane.
FT NON TER 1
SQ SEQUENCE 2087 AA; 222663 MW; 2C55D0554ECDD7A1 CRC64;

Query Match
Best Local Similarity 22.2%; Score 18; DB 2; Length 2087;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 442 CSSSSAASC 450

RESULT 29
Q9L981.VIBCH PRELIMINARY; PRT; 50 AA.
AC Q9L981;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORF1.
DE Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxId=666;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=569b;
RA MEDLINE=20476433; PubMed=11021935;
RA Clark C.A., Purine L., Kaewrakon P., Focareta T., Manning P.A.;
RT "The Vibrio cholerae O1 chromosomal integrin.";
RL Microbiology 146:2605-2612(2000).
DR EMBL; AF173596; AAF71194.1; -; Genomic DNA.
SQ SEQUENCE 50 AA; 5564 MW; 1A18FC4D3634E57 CRC64;

Query Match
Best Local Similarity 22.2%; Score 17; DB 2; Length 50;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 28 CASLASSSC 36

RESULT 30
Q4RA29.TETNG PRELIMINARY; PRT; 71 AA.
AC Q4RA29;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAP24619, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00037393001;
OC Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxId=99883;
RN [1]

```

RP NUCLEOTIDE SEQUENCE.  
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,  
RA Mauceli E., Bouneau L., Fischer C., Oufou-Costat C., Bernot A.,  
RA Nicud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segurens B.,  
RA Anthouard C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
RA Dandouard V., Jubin C., Castell V., Katinka M., Vacherie B.,  
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bernardinis V.,  
RA Cruaud C., Duprat S., Broctier P., Goutanceau J.P., Gouzy J.,  
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,  
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
RA Landet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,  
RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.;  
RT "Genome duplication in the teleost fish tetraodon nigroviridis reveals  
RT the early vertebrate proto-karyotype.";  
RL Nature 431:946-957(2004).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RG Genoscope, Whitehead Institute Centre for Genome Research;  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL: CAE01024619; CAG14754.1; -; Genomic\_DNA.  
FT NON\_TER 1  
SQ SEQUENCE 71 AA; 8036 MW; 736AF2E349B9AC50 CRC64;  
  
Query Match 54.8%; Score 17; DB 2; Length 71;  
Best Local Similarity 22.2%; Pred. No. 66;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 4 CXXXXXXXXC 12  
DB 27 CSTSRATC 35  
  
RESULT 31  
Q8VD96\_MESAU PRELIMINARY; PRT; 96 AA.  
ID Q8VD96;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE 10 kDa protein precursor.  
GN Name=CC10;  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Cricetinae; Mesocricetus.  
OX NCBI\_TaxID=10036;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC Gutierrez-Sagel R., Nieto A.;  
RT "Cloning and sequencing of a cDNA encoding hamster uteroglobin/clara  
RT cell 10 kDa protein.";  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U37041; ALA1349.1; -; Genomic\_DNA.  
DR HSSP; P17559; IUTR.  
DR SMR; Q8VD96; 20-96.  
DR GO; GO:0005496; F:steroid binding; IEA.  
DR InterPro; IPR006039; Utergl.  
DR InterPro; IPR003628; Uteroglobin\_sub.  
DR InterPro; IPR000329; Uteroglobin\_subf.  
DR InterPro; IPR006038; Uteroglobin\_subf.  
DR PANTHER; PTHR10136; Uteroglobin\_sub; 1.  
DR Pfam; PF01099; Uteroglobin; 1.  
DR PRINTS; PR00486; UTEROGLBIN.  
DR ProDom; PD012475; Uteroglbn\_sub; 1.  
DR SMART; SM00096; URG; 1.  
DR PROSITE; PS00404; UTEROGLBIN\_2; 1.  
FT SIGNAL 1 18 Potential.  
SQ SEQUENCE 1 18 Potential.

SQ SEQUENCE 96 AA; 10509 MW; 5EB9CBDD46143389 CRC64;  
  
Query Match 54.8%; Score 17; DB 2; Length 96;  
Best Local Similarity 22.2%; Pred. No. 74;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 4 CXXXXXXXXC 12  
DB 16 CSSASSDTC 24  
  
RESULT 32  
Q4P4H6\_USTMA PRELIMINARY; PRT; 108 AA.  
ID Q4P4H6;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=UM04987.1;  
OS Ustilago maydis 521.  
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;  
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.  
OX NCBI\_TaxID=237631;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=521;  
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,  
RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,  
RA Archach K., Armbruster J., Bachantsang P., Baldwin J., Barry A.,  
RA Bayot T., Blithstein B., Bloom T., Biye J., Boguslavsky L.,  
RA Borowsky M., Boukhalter B., Brunache A., Butler J., Calixte N.,  
RA Calvo S., Camarata J., Campo K., Chang J., Cheshtang Y., Citron M.,  
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,  
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,  
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,  
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,  
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,  
RA Ghitke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,  
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,  
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,  
RA Jaffe D., Jones C., Kamel M., Kamat A., Kamysseilis M., Karlsson E.,  
RA Kelle C., Kieu A., Klesner P., Kodira C., Kulbokas E., Labutti K.,  
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,  
RA Lindblad-Toh K., Liu X., Lokysang T., Lokysang Y., Lucien O.,  
RA Lui A., Ma L.U., Mabbitt R., MacDonald J., Maclean C., Major J.,  
RA Manning J., Maraballa R., Maru K., Matthews C., Mauceli E.,  
RA McCarthy M., McDonough S., McInee T., Meidrim J., Meneus L.,  
RA Mesirov J., Mihaliev A., Mihova T., Mikelsen T., Mienga V., Moru K.,  
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,  
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,  
RA Norbu N., O'donnell F., Okowo O., O'leary S., Omotosho B.,  
RA O'Neill K., Oseman S., Parker S., Perrin D., Phunkhang P., Pigani B.,  
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,  
RA Retta R., Richardson S., Rise C., Rodrigues J., Rogers J., Rogov P.,  
RA Rutman M., Schubach R., Seaman C., Settillalli S., Sharpe T.,  
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnuez C.,  
RA Spencer B., Stalker J., Strange-thomann N., Svaropoulou S.,  
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas D., Tchinga P.,  
RA Tenzing P., Teafaye S., Theodore J., Thoulungang Y., Topham K.,  
RA Towse S., Tsamila T., Tsomo N., Vallée D., Vaasiliev H.,  
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,  
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yedav S.,  
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,  
RA Zimmer A., Zody M., Lander E.;  
RT "The genome sequence of Ustilago maydis.";  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AACP0100184; EAK85760.1; -; Genomic\_DNA.  
DR Hypothetical protein.  
SQ SEQUENCE 108 AA; 11407 MW; E641B76DB8C16E7A CRC64;

Query Match 54.8%; Score 17; DB 2; Length 108;  
 Best Local Similarity 22.2%; Pred. No. 78;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12  
 Db 14 CTSANATC 22

## RESULT 33

Q9ZSP6 ARATH PRELIMINARY; PRT; 108 AA.  
 ID Q9ZSP6; AC Q9ZSP6;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Putative cell wall-plasma membrane disconnecting CLCT protein.  
 GN Name=AR1B;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN (1)

RP NUCLEOTIDE SEQUENCE.  
 RA Neuhboom L.W., van der Zaal E.J.;  
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.

RN NUCLEOTIDE SEQUENCE.  
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,  
 RA Nakajima M., Enju A., Kamita A., Narusaka M., Carninci P., Kawai J.,  
 RA Hayashizaki Y., Shinzaki K.;  
 RL Submitted (Nov-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF098631; AAD12259.1; -; genomic\_DNA.  
 DR HSSP: P24337; IHYP.  
 DR InterPro: IPR003612; AAI.  
 DR Pfam: PF00234; Tryp\_alpha\_amy1; 1.  
 DR SMART: SM00499; AAI; 1.194 MW; 8CD9F6CAE2577D2 CRC64;  
 SQ SEQUENCE 108 AA; 1194 MW; 8CD9F6CAE2577D2 CRC64;

Query Match 54.8%; Score 17; DB 2; Length 108;  
 Best Local Similarity 22.2%; Pred. No. 78;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12  
 Db 20 CTSATGTC 28

## RESULT 34

Q72D14 DESVH PRELIMINARY; PRT; 115 AA.  
 ID Q72D14; AC Q72D14;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN Ordered locus names=DVU117;  
 OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8103).  
 OC Bacteria; Proteobacteria; DeltaProteobacteria; Desulfovibrionales;  
 OC Desulfovibrionaceae; Desulfovibrio.  
 OC NCBI\_TaxID=882;  
 RN (1)

RP NUCLEOTIDE SEQUENCE.  
 RX PubMed=15077118; DOI=10.1038/nbt959;  
 RA Heideberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,  
 RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,  
 RA Daugherty S.C., Deboy R.T., Dodson R.U., Durkin A.S., Madupu R.,  
 RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,  
 RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,

RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Uterback T.R.,  
 RA Feldblum T.V., Wall J.D., Voordouw G., Fraser C.M.;  
 RT "The genome sequence of the anaerobic, sulfate-reducing bacterium  
 RT Desulfovibrio vulgaris Hildenborough.";  
 RL Nat. Biotechnol. 22:554-559(2004).  
 DR EMBL: AE017313; AAS95597.1; -; genomic\_DNA.  
 DR TIGR: DVU1117; -;  
 KM Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 115 AA; 11838 MW; 8E8A4DDF3EB96E64 CRC64;

Qy 4 CXXXXXXXXC 12  
 Db 70 CDTATASC 78

Query Match 54.8%; Score 17; DB 2; Length 115;  
 Best Local Similarity 22.2%; Pred. No. 80;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12  
 Db 70 CDTATASC 78

## RESULT 35

O6VZRI CNPV PRELIMINARY; PRT; 117 AA.  
 ID O6VZRI; AC O6VZRI;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE CNPV086 TNFR-like protein.  
 GN Name=CNPV086;  
 OS Canaripox virus (CNPV).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Avipoxvirus.  
 OC NCBI\_TaxID=44088;  
 RN (1)

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ATCC VR-111;  
 RX PubMed=14671117; DOI=10.1128/JVI.78.1.353-366.2004;  
 RA Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;  
 RT "The genome of canaripox virus.";  
 RL J. Virol. 78:353-366(2004).  
 DR EMBL: AY318871; AAR83432.1; -; genomic\_DNA.  
 DR GO: GO:0004872; Fireceptor activity; IEA.  
 DR InterPro: IPR006209; EGF-like.  
 DR InterPro: IPR001368; TNFR\_c6.  
 DR Pfam: PF00020; TNFR\_c6; 1.  
 DR SMART: SM00208; TNFR; 2.  
 DR PROSITE: PS01186; EGF 2; UNKNOWN\_1.  
 DR PROSITE: PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE: PS50050; TNFR\_NGFR\_2; 1.  
 SQ SEQUENCE 117 AA; 13113 MW; 06E99CAB6A41135A CRC64;

Query Match 54.8%; Score 17; DB 2; Length 117;  
 Best Local Similarity 22.2%; Pred. No. 80;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 12  
 Db 50 CTSSTTTC 58

## RESULT 36

Q4TWL2 ASF PRELIMINARY; PRT; 119 AA.  
 ID Q4TWL2; AC Q4TWL2;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE 9RL protein (Fragment).  
 GN Name=B602L;  
 OS African swine fever virus (ASFV).  
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfarvirus.  
 OC NCBI\_TaxID=10497;  
 RN (1)

RP NUCLEOTIDE SEQUENCE.  
 RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,



```

RC STRAIN=Portalegre 90;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
fever virus.";
RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028332; AAY46118.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 54.8%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 37
Q4TWL3 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL3_
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
fever virus.";
RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028331; AAY46117.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 54.8%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 38
Q4TWL5 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL5_
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
fever virus.";
RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028329; AAY46115.1; -; Genomic_DNA.
FT NON_TER 1

```

```

FT NON_TER 119
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 54.8%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 39
Q4TWL6 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL6_
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
fever virus.";
RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028328; AAY46114.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 119 AA; 11863 MW; 3A062A08CBA8D31F CRC64;

Query Match
Best Local Similarity 54.8%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12
DB 33 CASTGASTC 41

RESULT 40
Q4TWL7 ASF PRELIMINARY; PRT; 119 AA.
AC Q4TWL7_
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
fever virus.";
RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028327; AAY46113.1; -; Genomic_DNA.
FT NON_TER 1
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 54.8%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```



QY 4 CXXXXXXXXC 12  
DB 33 CASTGASTC 41

## RESULT 41

Q4TWL8 ASF PRELIMINARY; PRT; 119 AA.  
AC Q4TWL8;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DE 9RL protein (Fragment).  
GN Name=B602L;  
OS African swine fever virus (ASFV).  
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.  
OX NCBI\_TaxID=10497;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Almodovar 99;  
RA Duarte M.M., Cruz M.B.;  
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; DQ028326; AAY46112.1; -; Genomic\_DNA.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 54.8%; Score 17; DB 2; Length 119;  
Best Local Similarity 22.2%; Pred. No. 81;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 33 CASTGASTC 41

## RESULT 42

Q4TWL9 ASF PRELIMINARY; PRT; 119 AA.  
AC Q4TWL9;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE 9RL protein (Fragment).  
GN Name=B602L;  
OS African swine fever virus (ASFV).  
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.  
OX NCBI\_TaxID=10497;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Almodovar 99/NE1;  
RA Duarte M.M., Cruz M.B.;  
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; DQ028325; AAY46111.1; -; Genomic\_DNA.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 54.8%; Score 17; DB 2; Length 119;  
Best Local Similarity 22.2%; Pred. No. 81;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 33 CASTGASTC 41

## RESULT 43

Q4TWM0 ASF

ID Q4TWM0 ASF PRELIMINARY; PRT; 119 AA.

AC Q4TWM0;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DE 9RL protein (Fragment).  
GN Name=B602L;  
OS African swine fever virus (ASFV).  
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.  
OX NCBI\_TaxID=10497;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Almodovar 99/E2;  
RA Duarte M.M., Cruz M.B.;  
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; DQ028324; AAY46110.1; -; Genomic\_DNA.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 54.8%; Score 17; DB 2; Length 119;  
Best Local Similarity 22.2%; Pred. No. 81;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 33 CASTGASTC 41

## RESULT 44

Q5YPN5 NOCFA PRELIMINARY; PRT; 125 AA.  
ID Q5YPN5;  
AC Q5YPN5;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Hypothetical protein.  
GN OrderedlocusNames=nf550040;  
OS Nocardia farcinica.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Nocardiaceae; Nocardia.  
OX NCBI\_TaxID=37329;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=IFM 10152;  
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;  
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K., Shiba T., Hattori M.;  
RT "The complete genomic sequence of Nocardia farcinica IFM 10152."  
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).  
DR EMBL; AP006618; BAD59856.1; -; Genomic\_DNA.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 125 AA; 12749 MW; 1FD69A1CADF924EB CRC64;

Query Match 54.8%; Score 17; DB 2; Length 125;  
Best Local Similarity 22.2%; Pred. No. 83;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 12  
DB 96 CRSSSATTC 104

## RESULT 45

Q4T2U4 TETNG PRELIMINARY; PRT; 133 AA.  
ID Q4T2U4 TETNG;  
AC Q4T2U4;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Chromosome 12 SCAFI0174, whole genome shotgun sequence.

DE (Fragment).  
GN ORFNames=GSTENG0008171001;  
OS Tetraodon nigroviridis (Green puffer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
OC Tetraodontidae; Tetraodontidae; Tetraodon.  
OX NCB1\_TaxID=99883;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Coostat C., Bernot A.,  
RA Nicud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segures B.,  
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
RA Antouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,  
RA Bismont C., Skali Z., Cattolico L., Poulain J., De Berardinis V.,  
RA Craud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,  
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,  
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
RA Laudet V., Schachter V., Queller F., Saurin W., Scarpelli C.,  
RA Winkler P., Lander E.S., Weissbach J., Roest Crollius H.;  
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
RT the early vertebrate proto-karyotype.";  
RL Nature 431:946-957(2004).  
RN [2]  
RN NUCLEOTIDE SEQUENCE.  
RP GenomeScope; Whitehead Institute Centre for Genome Research;  
RG Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
RL -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; CAEO1010174; CAF92788.1; -; Genomic\_DNA.  
FT NON TER 1  
SQ SEQUENCE 133 AA; 14711 MW; 528890284B8084F4 CRC64;  
OY 4 CXXXXXXC 12  
DB 122 CGATSTTC 130

Query Match 54.8%; Score 17; DB 2; Length 133;  
Best Local Similarity 22.2%; Pred. No. 85;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 46  
O6ZRF5 HUMAN  
ID O6ZRF5 HUMAN PRELIMINARY; PRT; 134 AA.  
AC O6ZRF5;  
DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DE Hypothetical protein FLN46388.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;  
OC Homo.  
OX NCB1\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Thymus.  
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,  
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,  
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,  
RA Isono Y., Kawai-Hio Y., Saiko K., Nishikawa T., Kimura K.,  
RA Yamaehita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,  
RA Magatsuna M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,  
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nishigaki K.,  
RA Masuno Y., Nagai K., Isogai T.,  
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK128253; BAC87355.1; -; mRNA.  
SQ SEQUENCE 134 AA; 14481 MW; 03BA7E349C87BC6A CRC64;

Query Match 54.8%; Score 17; DB 2; Length 134;  
Best Local Similarity 22.2%; Pred. No. 85;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CXXXXXXC 12  
DB 27 CGASAAAC 35

RESULT 47  
O9CWC3 MOUSE  
ID O9CWC3 MOUSE PRELIMINARY; PRT; 147 AA.  
AC O9CWC3;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DE Mus musculus ES cells cDNA, RIKEN full-length enriched library,  
DE clone:2410101B22 product:LENG5 PROTEIN homolog.  
GN Name=Leng5;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCB1\_TaxID=10090;  
RN [1]  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=C57BL/6J;  
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RX Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=C57BL/6J;  
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RX Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Komoto H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saiko R.,  
RA Kadota K., Matuda H.A., Ashburner M., Batilov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Batsch G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.U., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=C57BL/6J;  
RC The FANTOM Consortium;  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [4]  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=C57BL/6J;  
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RX Carninci P., Shibata Y., Hayatsu N., Sugihara Y., Shibata K., Itoh M.,  
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
RT "Normalization and subcloning of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,  
 RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Mishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanishi M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-Format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J;  
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Tsumatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK010759; BAB27164.1; -; mRNA.  
 DR Ensembl; ENSMUSG0000035585; Mus musculus.  
 DR MGI; MGI:1913328; Leng5.  
 SQ SEQUENCE 147 AA; 15856 MW; 988434686F9EA74 CRC64;  
 Query Match 54.8%; Score 17; DB 2; Length 147;  
 Best Local Similarity 22.2%; Pred. No. 88;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 CXXXXXXX 12  
 Db 19 CATACTCTC 27

RESULT 48  
 Q6F340\_ORYSA PRELIMINARY; PRT; 155 AA.  
 AC Q6F340;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DE Hypothetical protein OSUNBA0088106.14.  
 GN Name=OSUNBA0088106.14;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoideae; Oryza.  
 OC NCBI\_TaxID=39947;  
 RX NCBI\_TaxID=39947;  
 RP NUCLEOTIDE SEQUENCE.  
 RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,  
 RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,  
 RA Chen Y.-L., Chang C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,  
 RA Heiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,  
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,  
 RA Wu H.-P., Shaw J.-F.;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC129718; AAT69648.1; -; Genomic\_DNA.  
 DR Gramene; Q6F340; -;  
 KM Hypothetical protein.  
 SQ SEQUENCE 155 AA; 16979 MW; 0EADP4C906C39295 CRC64;  
 Query Match 54.8%; Score 17; DB 2; Length 155;  
 Best Local Similarity 22.2%; Pred. No. 90;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXX 12  
 Db 146 CRASASTC 154

RESULT 49  
 Q4TWL4\_ASF PRELIMINARY; PRT; 155 AA.  
 AC Q4TWL4;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE 9RL protein (Fragment).  
 GN Name=B602L;  
 OS African swine fever virus (ASFV).  
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfarivirus.  
 OC NCBI\_TaxID=10497;  
 RX NCBI\_TaxID=10497;  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NH/pe8;  
 RA Duarte M.M., Cruz M.B.;  
 RT "B602L (9RL) partial gene of Portuguese isolates of African swine  
 RT fever virus.";  
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; DQ028330; AAY46116.1; -; Genomic\_DNA.  
 FT NON\_TER 1 1  
 FT NON\_TER 155 155  
 SQ SEQUENCE 155 AA; 15125 MW; 31AD66598783CE06 CRC64;  
 Query Match 54.8%; Score 17; DB 2; Length 155;  
 Best Local Similarity 22.2%; Pred. No. 90;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 CXXXXXXX 12  
 Db 69 CASTGASTC 77

RESULT 50  
 C1025\_HUMAN STANDARD; PRT; 157 AA.  
 ID C1025\_HUMAN  
 AC Q81W50; Q5T591; Q8T428;  
 DT 13-SEP-2005 (Rel. 48, Created)  
 DT 13-SEP-2005 (Rel. 48, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Protein C9orf25.  
 GN Name=C9orf25;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;  
 OC Homo.  
 OC NCBI\_TaxID=9606;  
 RX NCBI\_TaxID=9606;  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RX PubMed=15164053; DOI=10.1038/nature02465;  
 RA Humphrey S.J., Oliver K., Hunt A.R., Plumb R.W., Loveland J.E.,  
 RA Howe K.L., Andrews T.D., Searle S., Hunt S.E., Scott C.E., Jones M.C.,  
 RA Ainscough R., Almeida J.P., Ambrose K.D., Ashwell R.I.S.,  
 RA Babage A.K., Babage S., Baguley C.L., Bailey J., Banerjee R.,  
 RA Barker D.J., Barlow K.F., Bates K., Beasley H., Beasley O., Bird C.P.,  
 RA Bray-Allen S., Brown A.J., Brown J.Y., Burford D., Burrill W.,  
 RA Burton J., Carter C., Carter N.P., Chapman J.C., Chen Y., Clarke G.,  
 RA Clark S.Y., Clee C.M., Clegg S., Collier R.E., Corby N., Croxier M.,  
 RA Cummings A.T., Davies J., Dhami P., Dunn M., Dutta I., Dyer L.W.,  
 RA Earlrowl M.E., Faulkner L., Fleming C.J., Franklin A.,  
 RA Frankland J.A., French L., Fricker D.G., Garner P., Garnett J.,  
 RA Ghori J., Gilbert J.G.R., Glison C., Graham D.V., Gribble S.,  
 RA Griffiths C., Griffiths-Jones S., Grocock R., Guy J., Hall R.E.,  
 RA Hammond S., Harley J.L., Harrison E.S.I., Hart E.A., Heath P.D.,  
 RA Henderson C.D., Hopkins B.L., Howard P.J., Howden P.J., Huckle E.,  
 RA Johnson C., Johnson D., Joy A.A., Kay M., Keenan S., Kershaw J.K.,  
 RA Kimberley A.M., King A., Knights A., Laird G.K., Langford C.,

RA Lawlor S., Leongamornlert D.A., Leverisha M., Lloyd C., Lloyd D.M.,  
RA Lovell J., Martin S., Mashreghi-Mohammadi M., Matthews L., McLaren S.,  
RA McIlroy K.E., McMurray A., Milne S., Nickerson T., Nisbett J.,  
RA Nordiek G., Pearce A.V., Peck A.I., Porter K.M., Pandian R.,  
RA Pelan S., Phillimore B., Povey S.K., Ramey Y., Rand V., Scharfe M.,  
RA Sehra H.K., Showkeen R., Sims S.K., Stuce C.D., Smith M.,  
RA Steward C.A., Swarbreck D., Symore N., Tester J., Thorpe A.,  
RA Tracey A., Tromans A., Thomas D.W., Wall M., Wallis J.M., West A.P.,  
RA Whitehead S.L., Willey D.L., Williams S.A., Wilming L., Wray P.W.,  
RA Young L., Ashurst J.L., Coulson A., Blocker H., Durbin R.,  
RA Sulston J.E., Hubbard T., Jackson M.J., Bentley D.R., Beck S.,  
RA Rogers J., Dunham I.,  
RT "DNA sequence and analysis of human chromosome 9.";  
RL Nature 429:369-374 (2004).  
[2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,  
RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC EMBL; AL160270; CA113153.1; ALT INIT; Genomic DNA.  
DR EMBL; AL356494; CA113153.1; JOINED; Genomic DNA.  
DR EMBL; AL356494; CA115974.1; ALT INIT; Genomic DNA.  
DR EMBL; AL160270; CA115974.1; JOINED; Genomic DNA.  
DR EMBL; BC025348; AAH25348.1; -; mRNA.  
DR EMBL; BC041009; AAH41009.1; -; mRNA.  
DR Ensembl; ENSG00000164970; Homo sapiens.  
DR HGN; HGNC:19920; C9orf25.  
FT CONFLICT 106 106 Missing (in Ref. 2; AAH41009).  
SQ SEQUENCE 157 AA; 17365 MW; 577A4FA46B7DC74 CRC64;

Query Match 54.8%; Score 17; DB 1; Length 157;  
Best Local Similarity 22.2%; Pred. No. 90;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 CXXXXXXC 12  
DB 146 CQATSTAC 154

Search completed: January 4, 2006, 16:08:57  
Job time : 127.957 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 ; Search time 109.496 Seconds  
(without alignments)  
64.204 Million cell updates/sec

Title: US-09-932-322-4

Perfect score: 32

Sequence: 1 XXXXXXXXXXXXXXXX 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A\_Geneseq\_21:\*

1: geneeqp19808:\*  
2: geneeqp19908:\*  
3: geneeqp20008:\*  
4: geneeqp20018:\*  
5: geneeqp20028:\*  
6: geneeqp20038:\*  
7: geneeqp20048:\*  
8: geneeqp20058:\*  
9: geneeqp20068:\*

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	56.2	30	AD247471	Ad247471 Alu sense
2	18	56.2	89	ADP30533	Adp30533 Human sec
3	18	56.2	89	ADP30531	Adp30531 Human sec
4	18	56.2	109	AAW93527	AAW93527 Human pol
5	18	56.2	109	ADL31230	Adl31230 Human pro
6	18	56.2	109	ADP30690	Adp30690 Human sec
7	18	56.2	144	ADP31474	Adp31474 Human sec
8	18	56.2	167	ABO69151	ABO69151 Pseudomon
9	18	56.2	169	ADL312185	ADL312185 Human the
10	18	56.2	171	ADP30794	Adp30794 Human sec
11	18	56.2	171	ADP30793	Adp30793 Human sec
12	18	56.2	183	ADP30806	Adp30806 Human sec
13	18	56.2	191	ABO76619	ABO76619 Pseudomon
14	18	56.2	214	ADZ56213	ADZ56213 Human sec
15	18	56.2	228	ADP30921	Adp30921 Human sec
16	18	56.2	252	ADP31485	Adp31485 Human sec
17	18	56.2	258	ADP30479	Adp30479 Human sec
18	18	56.2	264	ADP31412	Adp31412 Human sec
19	18	56.2	270	ADP30500	Adp30500 Human sec
20	18	56.2	270	ADP31321	Adp31321 Human sec
21	18	56.2	294	ADP31473	Adp31473 Human sec
22	18	56.2	297	ADP31192	Adp31192 Human sec
23	18	56.2	328	AAW25285	AAW25285 Human pro
24	18	56.2	339	ADP30702	Adp30702 Human sec

25	18	56.2	348	8	ADP31441	Adp31441 Human sec
26	18	56.2	356	4	ABG21039	Abg21039 Novel hum
27	18	56.2	357	8	ADP31267	Adp31267 Human sec
28	18	56.2	357	8	ADP30505	Adp30505 Human sec
29	18	56.2	360	8	ADP31439	Adp31439 Human sec
30	18	56.2	390	8	ADP31218	Adp31218 Human sec
31	18	56.2	421	8	ADP31159	Adp31159 Human sec
32	18	56.2	426	8	ADP31495	Adp31495 Human sec
33	18	56.2	438	7	ABO82561	ABO82561 Pseudomon
34	18	56.2	453	8	ADP31465	Adp31465 Human sec
35	18	56.2	471	8	ADP30854	Adp30854 Human sec
36	18	56.2	525	8	ADP31227	Adp31227 Human sec
37	18	56.2	549	8	ADP30855	Adp30855 Human sec
38	18	56.2	555	8	ADP31416	Adp31416 Human sec
39	18	56.2	555	8	ADP31417	Adp31417 Human sec
40	18	56.2	588	8	ADP31699	Adp31699 Human sec
41	18	56.2	591	9	ADZ56209	ADZ56209 Human sec
42	18	56.2	591	9	ADZ56223	ADZ56223 Human sec
43	18	56.2	600	8	ADP30865	Adp30865 Human sec
44	18	56.2	604	8	ADP30940	Adp30940 Human sec
45	18	56.2	604	8	ADP30941	Adp30941 Human sec
46	18	56.2	605	8	ADP30507	Adp30507 Human sec
47	18	56.2	617	8	ADP31657	Adp31657 Human sec
48	18	56.2	626	6	ABR41596	ABR41596 Human DIT
49	18	56.2	645	8	ADP30858	Adp30858 Human sec
50	18	56.2	653	3	AAZ94907	AAZ94907 Human sec
51	18	56.2	653	3	ADZ73254	ADZ73254 Human kid
52	18	56.2	653	5	ADW42108	ADW42108 Protein e
53	18	56.2	653	7	ADH9077	ADH9077 TAT290. 1
54	18	56.2	653	8	ADL10948	ADL10948 Human the
55	18	56.2	653	9	ADZ56204	ADZ56204 Human KIA
56	18	56.2	653	9	ADZ56214	ADZ56214 Human KIA
57	18	56.2	660	8	ADP31677	Adp31677 Human sec
58	18	56.2	665	8	ADP30571	Adp30571 Human sec
59	18	56.2	681	8	ADP31582	Adp31582 Human sec
60	18	56.2	690	8	ADP30750	Adp30750 Human sec
61	18	56.2	708	8	ADP31191	Adp31191 Human sec
62	18	56.2	711	8	ADP31440	Adp31440 Human sec
63	18	56.2	728	8	ADP30508	Adp30508 Human sec
64	18	56.2	739	8	ADP31196	Adp31196 Human sec
65	18	56.2	750	8	ADP30576	Adp30576 Human sec
66	18	56.2	757	8	ADP30925	Adp30925 Human sec
67	18	56.2	762	8	ADP30891	Adp30891 Human sec
68	18	56.2	771	8	ADP31469	Adp31469 Human sec
69	18	56.2	772	8	ADP30936	Adp30936 Human sec
70	18	56.2	774	8	ADP30506	Adp30506 Human sec
71	18	56.2	774	8	ADP30780	Adp30780 Human sec
72	18	56.2	774	8	ADP30781	Adp30781 Human sec
73	18	56.2	795	8	ADP31332	Adp31332 Human sec
74	18	56.2	804	8	ADP31291	Adp31291 Human sec
75	18	56.2	813	8	ADP30649	Adp30649 Human sec
76	18	56.2	813	8	ADP30738	Adp30738 Human sec
77	18	56.2	852	8	ADP30664	Adp30664 Human sec
78	18	56.2	852	8	ADP30972	Adp30972 Human sec
79	18	56.2	865	8	ADP30566	Adp30566 Human sec
80	18	56.2	882	8	ADP31688	Adp31688 Human sec
81	18	56.2	882	8	ADP30971	Adp30971 Human sec
82	18	56.2	888	8	ADP31578	Adp31578 Human sec
83	18	56.2	892	8	ADP31444	Adp31444 Human sec
84	18	56.2	930	8	ADP31486	Adp31486 Human sec
85	18	56.2	947	8	ADP30937	Adp30937 Human sec
86	18	56.2	960	8	ADP31471	Adp31471 Human sec
87	18	56.2	960	8	ADP31470	Adp31470 Human sec
88	18	56.2	960	8	ADP31470	Adp31470 Human sec
89	18	56.2	1017	6	AAO26719	AAO26719 SR protei
90	18	56.2	1033	8	ADP30984	Adp30984 Human sec
91	18	56.2	1065	8	ADP30661	Adp30661 Human sec
92	18	56.2	1065	8	ADP30973	Adp30973 Human sec
93	18	56.2	1065	8	ADP31482	Adp31482 Human sec
94	18	56.2	1090	8	ADP310187	Adp310187 Human pol
95	18	56.2	1092	8	ADP31358	Adp31358 Human sec
96	18	56.2	1110	8	ADP31430	Adp31430 Human sec
97	18	56.2	1116	8	ADP31692	Adp31692 Human sec









536	17	53.1	1314	8	ADP31197	Human sec	609	17	53.1	1803	8	ADP30562	Adp30562	Human sec
537	17	53.1	1344	8	ADP31211	Human sec	610	17	53.1	1815	8	ADP31601	Adp31601	Human sec
538	17	53.1	1365	8	ADP31661	Human sec	611	17	53.1	1827	8	ADP31170	Adp31170	Human sec
539	17	53.1	1371	8	ADP31646	Human sec	612	17	53.1	1833	8	ADP30642	Adp30642	Human sec
540	17	53.1	1380	8	ADP31566	Human sec	613	17	53.1	1848	8	ADP31372	Adp31372	Human sec
541	17	53.1	1386	8	ADP31371	Human sec	614	17	53.1	1849	8	ADP30752	Adp30752	Human sec
542	17	53.1	1398	8	ADP31488	Human sec	615	17	53.1	1875	8	ADP31664	Adp31664	Human sec
543	17	53.1	1401	8	ADP30532	Human sec	616	17	53.1	1933	8	ADP30889	Adp30889	Human sec
544	17	53.1	1417	8	ADP31160	Human sec	617	17	53.1	1933	8	ADP30902	Adp30902	Human sec
545	17	53.1	1417	8	ADP31357	Human sec	618	17	53.1	1956	8	ADP31662	Adp31662	Human sec
546	17	53.1	1454	8	ADP31177	Human sec	619	17	53.1	1968	8	ADP30689	Adp30689	Human sec
547	17	53.1	1456	8	ADP30923	Human sec	620	17	53.1	2001	8	ADP31644	Adp31644	Human sec
548	17	53.1	1464	8	ADP31040	Human sec	621	17	53.1	2020	8	ADP31056	Adp31056	Human sec
549	17	53.1	1464	8	ADP31437	Human sec	622	17	53.1	2027	8	ADP31058	Adp31058	Human sec
550	17	53.1	1485	8	ADP31383	Human sec	623	17	53.1	2052	9	AE887635	AE887635	Human ino
551	17	53.1	1485	8	ADP31384	Human sec	624	17	53.1	2058	8	ADP31630	Adp31630	Human sec
552	17	53.1	1488	8	ADP31385	Human sec	625	17	53.1	2123	8	ADP30657	Adp30657	Human sec
553	17	53.1	1488	8	ADP31386	Human sec	626	17	53.1	2123	8	ADP30564	Adp30564	Human sec
554	17	53.1	1494	8	ADP31650	Human sec	627	17	53.1	2127	8	ADP31327	Adp31327	Human sec
555	17	53.1	1506	8	ADP30596	Human sec	628	17	53.1	2148	8	ADP30974	Adp30974	Human sec
556	17	53.1	1521	8	ADP30542	Human sec	629	17	53.1	2260	8	ADP30687	Adp30687	Human sec
557	17	53.1	1521	8	ADP30537	Human sec	630	17	53.1	2272	8	ADP31136	Adp31136	Human sec
558	17	53.1	1530	8	ADP31536	Human sec	631	17	53.1	2349	8	ADP30959	Adp30959	Human sec
559	17	53.1	1536	8	ADP31302	Human sec	632	17	53.1	2382	8	ADP31341	Adp31341	Human sec
560	17	53.1	1539	8	ADP31201	Human sec	633	17	53.1	2454	8	ADP30469	Adp30469	Human sec
561	17	53.1	1539	8	ADP31200	Human sec	634	17	53.1	2484	8	ADP66690	ADP66690	Human mls
562	17	53.1	1560	8	ADP31570	Human sec	635	17	53.1	2508	6	ADA15721	ADA15721	C. elegans
563	17	53.1	1560	8	ADP31139	Human sec	636	17	53.1	2535	8	ADP31146	Adp31146	Human sec
564	17	53.1	1560	8	ADP31631	Human sec	637	17	53.1	2544	6	ADA15717	ADA15717	C. elegans
565	17	53.1	1566	8	ADP31407	Human sec	638	17	53.1	2547	8	ADP31665	Adp31665	Human sec
566	17	53.1	1584	8	ADP31405	Human sec	639	17	53.1	2601	6	ADA15723	ADA15723	C. elegans
567	17	53.1	1587	8	ADP30591	Human sec	640	17	53.1	2622	8	ADP31663	Adp31663	Human sec
568	17	53.1	1596	8	ADP31491	Human sec	641	17	53.1	2664	8	ADP31546	Adp31546	Human sec
569	17	53.1	1615	4	AA859826	Protein #	642	17	53.1	2688	8	ADN11593	ADN11593	Human CD9
570	17	53.1	1623	8	ADP30552	Human sec	643	17	53.1	2724	8	ADP31233	Adp31233	Human sec
571	17	53.1	1631	4	ABG22481	Human sec	644	17	53.1	2790	8	ADP31232	Adp31232	Human sec
572	17	53.1	1647	8	ADP30670	Human sec	645	17	53.1	2808	8	ADP31258	Adp31258	Human sec
573	17	53.1	1647	8	ADP31052	Human sec	646	17	53.1	2827	8	ADP30681	Adp30681	Human sec
574	17	53.1	1652	6	ADA15715	C. elegans	647	17	53.1	2828	8	ADP30938	Adp30938	Human sec
575	17	53.1	1654	8	ADP30997	Human sec	648	17	53.1	2833	8	ADP31299	Adp31299	Human sec
576	17	53.1	1662	8	ADP31419	Human sec	649	17	53.1	2835	8	ADP30572	Adp30572	Human sec
577	17	53.1	1662	8	ADP31513	Human sec	650	17	53.1	3046	8	ADQ10419	ADQ10419	Human pol
578	17	53.1	1665	8	ADP31187	Human sec	651	17	53.1	3144	8	ADP31544	Adp31544	Human sec
579	17	53.1	1700	3	AA818144	Plasmodiu	652	17	53.1	3201	8	ADP31545	Adp31545	Human sec
580	17	53.1	1725	8	ADP30654	Human sec	653	17	53.1	3398	9	AE887634	AE887634	Human ino
581	17	53.1	1737	8	ADP31282	Human sec	654	17	53.1	3407	8	ADP31060	Adp31060	Human sec
582	17	53.1	1743	6	ABU88225	Human sec	655	17	53.1	3407	8	ADP31062	Adp31062	Human sec
583	17	53.1	1743	6	ABU90134	Human sec	656	17	53.1	3465	8	ADP31234	Adp31234	Human sec
584	17	53.1	1743	6	ABU96436	Human sec	657	17	53.1	3579	8	ADP31098	Adp31098	Human sec
585	17	53.1	1743	6	ABU99045	Human sec	658	17	53.1	3638	8	ADP30981	Adp30981	Human sec
586	17	53.1	1743	6	ABU98260	Human sec	659	17	53.1	4683	8	ADP31260	Adp31260	Human sec
587	17	53.1	1743	6	ABU91966	Human sec	660	17	53.1	4848	8	ADP31259	Adp31259	Human sec
588	17	53.1	1743	6	ABU85270	Human sec	661	17	53.1	5397	8	ADP31068	Adp31068	Human sec
589	17	53.1	1743	6	ABO00409	Human sec	662	17	53.1	5514	8	ADP31591	Adp31591	Human sec
590	17	53.1	1743	6	ABU88960	Human sec	663	17	53.1	5820	8	ADP31118	Adp31118	Human sec
591	17	53.1	1743	6	ABO06456	Human sec	664	17	53.1	5820	8	ADP31119	Adp31119	Human sec
592	17	53.1	1743	6	ABU95516	Human sec	665	17	53.1	8973	8	ADP31145	Adp31145	Human sec
593	17	53.1	1743	6	ABU95206	Human sec	666	17	53.1	8976	8	ADP31425	Adp31425	Human sec
594	17	53.1	1743	6	ABU90754	Human sec	667	17	53.1	9195	8	ADP31494	Adp31494	Human sec
595	17	53.1	1743	6	ABU93916	Human sec	668	17	53.1	10	2	AA871548	AA871548	Human sec
596	17	53.1	1743	6	ABU86190	Human sec	669	17	53.1	10	2	AA828918	AA828918	Human sec
597	17	53.1	1743	6	ABU82045	Human sec	670	17	53.1	22	5	ABO05791	ABO05791	Human sec
598	17	53.1	1743	6	ABU07906	Human sec	671	17	53.1	22	5	ADV76485	ADV76485	Human HER
599	17	53.1	1743	6	ABU94226	Human sec	672	17	53.1	25	5	AAE21859	AAE21859	Human HER
600	17	53.1	1743	6	ABO00099	Human sec	673	17	53.1	25	5	AAE19915	AAE19915	Human HER
601	17	53.1	1743	6	ABU87110	Human sec	674	17	53.1	25	7	ABW00366	ABW00366	Human HER
602	17	53.1	1743	6	ABU91351	Human sec	675	17	53.1	25	8	ADG47681	ADG47681	Human HER
603	17	53.1	1743	6	ABU90444	Human sec	676	17	53.1	34	5	AAU87022	AAU87022	Human HER
604	17	53.1	1743	6	ABU97035	Human sec	677	17	53.1	41	4	ABBA1607	ABBA1607	Human HER
605	17	53.1	1743	6	ABO05231	Human sec	678	17	53.1	41	4	AAW35400	AAW35400	Human HER
606	17	53.1	1749	8	ADP31408	Human sec	679	17	53.1	41	4	ABR25433	ABR25433	Human HER
607	17	53.1	1755	8	ADP31446	Human sec	680	17	53.1	41	4	AAW75287	AAW75287	Human HER
608	17	53.1	1789	8	ADP30962	Human sec	681	17	53.1	41	4	AAW62478	AAW62478	Human bta

682	16	50.0	41	4	ABG57048	Abg57048	Human 11v	755	110	6	ABM65599	Abm65599	Propionib
683	16	50.0	41	5	ABG44927	Abg44927	Human pep	756	111	5	ADK35517	Adk35517	Novel hum
684	16	50.0	44	2	AAR98208	Aar98208	Nucleotid	757	112	7	ABO68317	AbO68317	Pseudomon
685	16	50.0	45	8	ADT92777	Adt92777	Mouse Cal	758	112	8	ADX88804	Adx88804	Plant ful
686	16	50.0	47	8	ADT92788	Adt92788	Mouse Cal	759	113	8	ABY03134	AbY03134	SARS coro
687	16	50.0	45	8	ADZ47821	Adz47821	Amino aci	760	113	9	AEB91569	Aeb91569	Microbial
688	16	50.0	48	3	AAGS7161	Aags7161	Arabiidops	761	114	5	ABP06472	Abp06472	Human ORF
689	16	50.0	48	3	AAAG60935	Aag60935	Arabiidops	762	115	4	AAH83845	Aah83845	Human imm
690	16	50.0	48	3	ADV76464	Adv76464	Human HER	763	116	8	ADP30955	Adp30955	Human sec
691	16	50.0	48	9	AEA34233	Aea34233	Optum pop	764	118	4	ABG22738	Abg22738	Novel hum
692	16	50.0	51	8	ADP29386	Adp29386	Human sec	765	121	9	ABM96843	Abm96843	M. xanthu
693	16	50.0	51	8	ADP29383	Adp29383	Human sec	766	122	4	ABH17561	Abh17561	Human ner
694	16	50.0	52	5	ABP31338	Abp31338	Human ORF	767	122	7	ABO83310	AbO83310	Pseudomon
695	16	50.0	52	5	AAU48534	Aau48534	Propionib	768	123	4	AAH90307	Aah90307	Human imm
696	16	50.0	58	6	ABM45053	Abm45053	Propionib	769	125	4	ABG04861	Abg04861	Novel hum
697	16	50.0	61	4	AAH19250	Aah19250	Peptide #	770	133	4	AAU64082	Aau64082	Propionib
698	16	50.0	61	4	ABH38496	Abh38496	Peptide #	771	133	6	ABM60601	Abm60601	Propionib
699	16	50.0	61	4	AAH31941	Aah31941	Peptide #	772	133	8	ADP31069	Adp31069	Human sec
700	16	50.0	61	4	ABH23648	Abh23648	Protein #	773	134	7	ABO75321	AbO75321	Human sec
701	16	50.0	61	4	AAH71647	Aah71647	Human bon	774	136	7	ABO73061	AbO73061	Pseudomon
702	16	50.0	61	4	AAH59112	Aah59112	Human bra	775	137	4	AAH93766	Aah93766	Human pro
703	16	50.0	61	4	ABG53331	Abg53331	Human 11v	776	137	7	ABO82722	AbO82722	Pseudomon
704	16	50.0	61	5	ABG41461	Abg41461	Human pep	777	139	7	ABO80767	AbO80767	Pseudomon
705	16	50.0	63	4	AAU59447	Aau59447	Propionib	778	139	7	ABO72721	AbO72721	Pseudomon
706	16	50.0	63	6	ABM55966	Abm55966	Propionib	779	139	7	ABO76095	AbO76095	Pseudomon
707	16	50.0	63	6	ABR63898	AbR63898	P. furiosus	780	140	7	ABO83055	AbO83055	Pseudomon
708	16	50.0	65	4	ABP35884	Abp35884	CYR1-GADH	781	140	7	ABO83386	AbO83386	Pseudomon
709	16	50.0	65	4	ABP35883	Abp35883	CYR1-GADH	782	141	4	AAH63384	Aah63384	Human pro
710	16	50.0	65	4	AAU41244	Aau41244	Propionib	783	146	7	ABO84154	AbO84154	Pseudomon
711	16	50.0	65	6	ABM37763	Abm37763	Propionib	784	146	7	ABO72664	AbO72664	Pseudomon
712	16	50.0	66	3	AAH57014	Aah57014	Arabiidops	785	147	4	ABG29376	AbG29376	Novel hum
713	16	50.0	68	3	ADP31694	Adp31694	Human sec	786	147	7	ABO79762	AbO79762	Pseudomon
714	16	50.0	70	7	ABO73323	AbO73323	Pseudomon	787	150	2	AAH49809	Aah49809	Human gli
715	16	50.0	71	4	AAH41601	Aah41601	Peptide #	788	150	3	AAH20620	Aah20620	Novel hum
716	16	50.0	71	4	ABH33561	Abh33561	Peptide #	789	150	7	ABO64648	AbO64648	Klebsiell
717	16	50.0	71	4	AAH27022	Aah27022	Peptide #	790	152	8	ADP31605	Adp31605	Human sec
718	16	50.0	71	4	ABH28383	Abh28383	Peptide #	791	153	8	ADY95514	AdY95514	Human SAR
719	16	50.0	71	4	ABH19018	Abh19018	Protein #	792	154	5	ABH62822	AbH62822	Human pol
720	16	50.0	71	4	AAH66735	Aah66735	Human bon	793	154	6	AAH36181	Aah36181	Human CBD
721	16	50.0	71	4	AAU53020	Aau53020	Propionib	794	155	7	ABO77012	AbO77012	Pseudomon
722	16	50.0	71	4	AAH54335	Aah54335	Human bra	795	157	7	ABO83103	AbO83103	Pseudomon
723	16	50.0	71	4	ABH48403	Abh48403	Human 11v	796	158	7	ABO76328	AbO76328	Pseudomon
724	16	50.0	71	4	AAH02336	Aah02336	Peptide #	797	158	7	ABO79744	AbO79744	Pseudomon
725	16	50.0	71	5	ABH36389	Abh36389	Human pep	798	158	8	ABO59415	AbO59415	Human gen
726	16	50.0	71	6	ABH49533	Abh49533	Propionib	799	159	4	AAH010454	Aah010454	Human pol
727	16	50.0	71	9	ADZ47683	Adz47683	Anti-IL-1	800	159	7	ABO79796	AbO79796	Pseudomon
728	16	50.0	74	4	AAU47686	Aau47686	Propionib	801	160	7	ABO73965	AbO73965	Pseudomon
729	16	50.0	74	6	ABH44205	Abh44205	Propionib	802	162	3	AAH95932	Aah95932	Porcine a
730	16	50.0	80	4	AAH70800	Aah70800	Human bon	803	162	7	ABO69461	AbO69461	Pseudomon
731	16	50.0	80	4	AAU59526	Aau59526	Propionib	804	162	7	ABO76207	AbO76207	Pseudomon
732	16	50.0	80	6	ABH56045	Abh56045	Propionib	805	162	7	ABO79469	AbO79469	Pseudomon
733	16	50.0	88	4	AAH74029	Aah74029	Human col	806	162	8	ADP30813	Adp30813	Human sec
734	16	50.0	90	8	ADP31399	Adp31399	Human sec	807	162	8	ADH88496	AdH88496	Plant ful
735	16	50.0	90	8	ADP31400	Adp31400	Human sec	808	165	4	AAU54621	Aau54621	Propionib
736	16	50.0	92	4	AAU46180	Aau46180	Human sec	809	165	6	ABH51140	AbH51140	Propionib
737	16	50.0	92	6	ABH42699	Abh42699	Propionib	810	165	8	ADP31286	Adp31286	Human sec
738	16	50.0	93	4	ABH43737	Abh43737	Peptide #	811	165	8	ADP31174	Adp31174	Human sec
739	16	50.0	93	4	AAH37635	Aah37635	Peptide #	812	167	2	AAH48969	Aah48969	Neisseria
740	16	50.0	93	4	AAH77475	Aah77475	Human bon	813	167	9	AEH48987	Aeh48987	N. mening
741	16	50.0	93	5	ABH46494	Abh46494	Human pep	814	168	7	ABO80868	AbO80868	Pseudomon
742	16	50.0	93	8	ADP30859	Adp30859	Human sec	815	168	8	ADP30841	Adp30841	Human sec
743	16	50.0	95	8	ADG22362	Adg22362	Cyanophag	816	171	8	ABO60389	AbO60389	Human gen
744	16	50.0	104	3	AAH12843	Aah12843	HTLV-1 pr	817	171	8	ADP30599	Adp30599	Human sec
745	16	50.0	104	3	AAH24831	Aah24831	Arabiidops	818	173	7	ABO74910	AbO74910	Pseudomon
746	16	50.0	104	6	AAU5097	Aau5097	Propionib	819	174	8	ADP30801	Adp30801	Human sec
747	16	50.0	104	6	ABH51616	Abh51616	Propionib	820	175	4	ABH68003	AbH68003	Drosophil
748	16	50.0	104	7	ABO78018	AbO78018	Pseudomon	821	180	8	ADP30820	Adp30820	Human sec
749	16	50.0	105	4	AAU42091	Aau42091	Propionib	822	180	8	ADP30825	Adp30825	Human sec
750	16	50.0	105	6	ABH38610	Abh38610	Propionib	823	180	8	ADP30828	Adp30828	Human sec
751	16	50.0	106	6	AAH58682	Aah58682	Propionib	824	180	8	ADP30821	Adp30821	Human sec
752	16	50.0	106	6	ABH55201	Abh55201	Propionib	825	180	8	ADP30826	Adp30826	Human sec
753	16	50.0	108	4	AAU42243	Aau42243	Propionib	826	180	8	ADP30827	Adp30827	Human sec
754	16	50.0	108	6	ABH38762	Abh38762	Propionib	827	181	8	ADP31029	Adp31029	Human sec



```
974 16 50.0 366 8 ADP31075 Adp31075 Human sec
975 16 50.0 369 7 ABO72044 Abo72044 Pseudomon
976 16 50.0 369 8 ADP31289 Adp31289 Human sec
977 16 50.0 372 8 ADP31610 Adp31610 Human sec
978 16 50.0 375 9 ADP06695 Adp06695 Cyclin-de
979 16 50.0 380 3 AAG04919 Aag04919 Arabidops
980 16 50.0 380 3 AAG32553 Aag32553 Arabidops
981 16 50.0 381 2 AAY38570 Aay38570 Neisseria
982 16 50.0 381 2 AAY75046 Aay75046 Neisseria
983 16 50.0 381 9 AEB48989 Aeb48989 N. mening
984 16 50.0 382 2 AAY38571 Aay38571 Neisseria
985 16 50.0 382 3 AAY75044 Aay75044 Neisseria
986 16 50.0 382 4 ABB70181 Abb70181 Drosophil
987 16 50.0 382 9 AEB48990 Aeb48990 N. gonorr
988 16 50.0 384 8 ADP31338 Adp31338 Human sec
989 16 50.0 384 8 ADP31318 Adp31318 Human sec
990 16 50.0 387 8 ADP31158 Adp31158 Human sec
991 16 50.0 387 9 ADZ47819 Adz47819 Antino act
992 16 50.0 389 8 ADX78458 Adx78458 Plant ful
993 16 50.0 389 8 ADY24278 Ady24278 Plant ful
994 16 50.0 391 7 ADN31030 Adn31030 Fruit fly
995 16 50.0 396 8 ADP31625 Adp31625 Human sec
996 16 50.0 396 8 ADP31378 Adp31378 Human sec
997 16 50.0 398 4 AAM40928 Aam40928 Human pol
998 16 50.0 398 7 ADE60215 Adf60215 Human con
999 16 50.0 398 7 ABO71161 Abo71161 Pseudomon
1000 16 50.0 398 8 ADP31604 Adp31604 Human sec
```

## ALIGNMENTS

```
RESULT 1
ID AD247471 standard; peptide; 30 AA.
XX AD247471:
AC
XX 30-JUN-2005 (first entry)
XX
DE Alu sense peptide SEQ ID NO 50.
XX
KM hyaluronan synthase; HAS; hyaluronidase; HYAL; HAS1; HAS2; HAS3; HYAL1;
KM HYAL2; HAL3; PH-20; cancer; hyperproliferation; inflammation; anabolic;
KM hypertensive; gynaecological; neuroprotective; antiplatelet; cyclostatic;
KM anti-inflammatory; endocrine-gen; immunosuppressive; gene therapy.
XX
OS Unidentified.
XX
PN WO2005035548-A1.
XX
PD 21-APR-2005.
XX
PF 11-OCT-2004; 2004WO-AU001383.
XX
PR 10-OCT-2003; 2003AU-00905551.
PR 01-DEC-2003; 2003AU-00906658.
XX
PA (MEDI-) MEDITECH RES LTD.
XX
PI Brown TJ, Brownlee GR;
XX
DR WPI; 2005-315540/32.
XX
PT Novel compound capable of reducing level, function or activity of
PT hyaluronan synthase or hyaluronidase, useful for treating cancer,
PT hyperproliferative condition, A-beta-lipoproteinemia, A-V, A beta-2-
PT microglobulin amyloidosis.
XX
PS Example 22; SEQ ID NO 50; 277bp; English.
XX
CC This invention describes a novel nucleic acid molecule capable of
CC reducing the level of hyaluronan synthase (HAS) or hyaluronidase (HYAL)
```

```
CC or the function or activity of HAS or HYAL. HAS is selected from HAS1,
CC HAS2 and HAS3. HYAL is chosen from HYAL1, HYAL2, HAL3 and PH-20. The
CC nucleic acid molecule is an oligonucleotide or its chemically modified
CC form comprising a chemically modified backbone or a non-natural
CC internucleoside linkage. The compound is an interactive molecule capable
CC of binding or otherwise associating with HAS and/or HYAL to reduce HAS
CC and/or HYAL function or activity. The novel molecule is useful for
CC treating cancer, hyperproliferative conditions or inflammatory
CC conditions. The products of the invention are useful for treating or
CC prophylaxis of a condition in subject, which involves administering to
CC the subject, an HAS and/or HYAL level-reducing or HAS and/or HYAL
CC activity reducing effective amount of active molecule. The products of
CC the invention have anabolic, hypertensive, gynaecological,
CC neuroprotective, antiplatelet, cyclostatic, anti-inflammatory, endocrine-
CC gen, and immunosuppressive activity. This sequence represents a peptide
CC used in the method of the invention. NOTE: The specification describes
CC SEQ ID NO 27-51 as nucleotide primers however the sequence listing
CC represents the nucleotides in a three letter amino acid code. The
CC nucleotide sequences have been made and are represented in { }.
XX
SQ Sequence 30 AA;
XX
Query Match 56.2%; Score 18; DB 9; Length 30;
Best local Similarity 20.0%; Pred. No. 2.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 13
DB 18 CTTAAATATAC 27
RESULT 2
ADP30533
ID ADP30533 standard; protein; 89 AA.
XX
AC ADP30533;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1300.
XX
KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR 29-AUG-2002; 2002US-0406616P.
PR 29-AUG-2002; 2002US-0406640P.
PR 29-AUG-2002; 2002US-0406642P.
PR 29-AUG-2002; 2002US-0406646P.
PR 29-AUG-2002; 2002US-0406653P.
PR 29-AUG-2002; 2002US-0406655P.
PR 29-AUG-2002; 2002US-0406656P.
PR 17-SEP-2002; 2002US-0410946P.
PR 17-SEP-2002; 2002US-0410947P.
PR 17-SEP-2002; 2002US-0410948P.
PR 17-SEP-2002; 2002US-0410949P.
PR 17-SEP-2002; 2002US-0410953P.
PR 17-SEP-2002; 2002US-0410957P.
PR 17-SEP-2002; 2002US-0410958P.
```

17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX  
DR WPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 2531; 428bp; English.  
XX  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
XX  
SQ Sequence 89 AA;

Query Match 56.2%; Score 18; DB 8; Length 89;  
Best Local Similarity 20.0%; Pred. No. 3.1;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXX 13  
Db 24 CTTAAAAAAC 33  
RESULT 3  
ADP30531  
ID ADP30531 standard; protein; 89 AA.  
XX  
AC ADP30531;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1298.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 29-AUG-2002; 2002US-0406667P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0410963P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411101P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471336P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486466P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 XX  
 XX  
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 DR WPI; 2004-348438/32.  
 XX  
 XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX  
 XX Claim 1; SEQ ID NO 2529; 428pp; English.  
 PS  
 XX  
 CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMB and is not in the specification.  
 CC  
 XX  
 SQ Sequence 89 AA;  
 Query Match 56.2%; Score 18; DB 8; Length 89;  
 Best Local Similarity 20.0%; Pred. No. 3.1;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 4 CXXXXXXXXX 13  
 DB 24 CTTAAAAAAC 33  
 ID  
 ID AAM93527 standard; protein; 109 AA.  
 XX  
 XX AAM93527;  
 AC  
 XX  
 XX 06-NOV-2001 (first entry)  
 DT  
 XX  
 DE Human polypeptide, SEQ ID NO: 3263.  
 XX  
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.  
 XX  
 OS Homo sapiens.

XX  
 PN EPI130094-A2.  
 XX  
 PD 05-SEP-2001.  
 XX  
 PF 07-JUL-2000; 2000EP-00114089.  
 XX  
 PR 08-JUL-1999; 99JP-00194486.  
 PR 11-JAN-2000; 2000JP-00118774.  
 PR 02-MAY-2000; 2000JP-00183765.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 XX  
 DR WPI; 2001-524255/58.  
 DR N-PSDB; AAK94457.  
 XX  
 PT 830 Primers useful for synthesizing full length cDNA clones and their use  
 PT in genetic manipulation.  
 XX  
 PS Claim 8; SEQ ID NO 3263; 1380pp + Sequence listing; English.  
 XX  
 CC The invention relates to primers for synthesizing full length cDNA  
 CC clones. 830 cDNA molecules encoding a human protein have been isolated  
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have  
 CC been determined. Primers for synthesizing the full length cDNA are useful  
 CC for clarifying the function of the protein encoded by the cDNA. The full  
 CC length clones were obtained by construction of full length enriched cDNA  
 CC libraries that were synthesised by the oligo-capping method. The primers  
 CC enable the production of the full length cDNA easily without any special  
 CC methods. The present sequence is a polypeptide encoded by a full length  
 CC human cDNA of the invention. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in CD-ROM  
 CC format directly from EPO  
 CC  
 XX  
 SQ Sequence 109 AA;  
 Query Match 56.2%; Score 18; DB 4; Length 109;  
 Best Local Similarity 20.0%; Pred. No. 3.2;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 4 CXXXXXXXXX 13  
 DB 4 CSSSATSTSC 13  
 ID  
 ID ADL31230 standard; protein; 109 AA.  
 XX  
 AC ADL31230;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Human protein encoded by a full length cDNA clone SegID 3263.  
 XX  
 KW human; medicine; signal transduction; glycoprotein; transcription;  
 KW oligo-capping method.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EPI396543-A2.  
 PN  
 PD 10-MAR-2004.  
 XX  
 PF 07-JUL-2000; 2003EP-00025638.  
 XX  
 PR 08-JUL-1999; 99JP-00194486.  
 PR 11-JAN-2000; 2000JP-00118774.  
 PR 02-MAY-2000; 2000JP-00183865.  
 PR 07-JUL-2000; 2000EP-00114089.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H,  
XX WPI; 2004-204755/20.  
DR N-PSDB; ADL31229.  
XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full  
PT length human cDNAs.  
XX Example 1; SEQ ID NO 3263; 1340bp; English.  
XX This invention relates to a novel primers useful for synthesizing full  
CC length cDNA molecules that encode human proteins. Specifically, it refers  
CC to secretory or membrane proteins that are potential therapeutic agents/  
CC target molecules in the field of medicine, and in particular genes  
CC encoding proteins that are associated with signal transduction,  
CC glycoproteins and transcription. The present invention describes a method  
CC for efficiently cloning a full length human cDNA from both the 5' and 3'  
CC ends using the oligo-capping method. This polypeptide sequence is a full  
CC length human protein of the invention.  
XX Sequence 109 AA;  
SQ  
Query Match 56.2%; Score 18; DB 8; Length 109;  
Best Local Similarity 20.0%; Pred. No. 3.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXC 13  
Db 4 CXXXXXXXC 13  
ADP30690 standard; protein; 109 AA.  
ID ADP30690 standard; protein; 109 AA.  
XX ADP30690;  
AC  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1457.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003MO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406656P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 02-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485233P.  
PR 08-JUL-2003; 2003US-0485244P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486909P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LR, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T,  
PI Pierre K, Wang Y, Wong JCB, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2688; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is

CC available on WIPWEB and is not in the specification.

XX Sequence 109 AA;

Query Match 56.2%; Score 18; DB 8; Length 109;  
Best Local Similarity 20.0%; Pred. No. 3.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXC 13  
|  
Db 96 CTTTATATC 105

RESULT 7  
ADP31474  
ID ADP31474 standard; protein; 144 AA.

XX ADP31474;

DT 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #2241.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN W02004035732-A2.

XX W02004035732-A2.

PD 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406640P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410961P.

PR 17-SEP-2002; 2002US-0411019P.

PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.

PR 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.

PR 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493373P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX PA  
XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kotakota S, Halshan L, Linemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX DR  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX CC  
XX Claim 1; SEQ ID NO 3472; 428bp; English.  
XX PS  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX SQ  
XX Sequence 144 AA;  
XX  
XX Query Match 56.2%; Score 18; DB 8; Length 144;  
XX Best Local Similarity 20.0%; Pred. No. 3.4;  
XX Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
XX  
XX QY 4 CXXXXXXXC 13  
XX |  
XX Db 72 CAATAAATC 81  
XX  
XX  
XX RESULT 8  
XX ABO69151  
XX ID ABO69151 standard; protein; 167 AA.  
XX  
XX ABO69151;  
XX AC  
XX 29-JUL-2004 (first entry)  
XX DT  
XX Pseudomonas aeruginosa polypeptide #1326.  
XX DB





PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485225P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493373P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;  
PI Pierre K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic.

PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2792; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 171 AA;  
Query Match 56.2%; Score 18; DB 8; Length 171;  
Best Local Similarity 20.0%; Pred. No. 3.5;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXX 13  
DB 146 CAATTAATAC 155  
RESULT 11  
ADP30793  
ID ADP30793 standard; protein, 171 AA.  
XX  
XX ADP30793;  
AC 12-AUG-2004 (first entry)  
DT  
XX  
DE Human secreted protein SEQ ID #1560.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
PN WO2004035732-A2.  
XX  
XX 29-APR-2004.  
PD  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.

17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411010P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 09-JUN-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485225P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493533P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Hsieh L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 2791; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
XX  
SQ Sequence 171 AA;

Query Match 56.2%; Score 18; DB 8; Length 171;

Best Local Similarity 20.0%; Pred. No. 3.5;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 13  
DB 146 CAAATAAAC 155

RESULT 12  
ADP30806  
ID ADP30806 standard; protein; 183 AA.  
XX  
XX ADP30806;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1573.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411043P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411010P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 17-SEP-2002; 2002US-0463700P.  
XX 18-APR-2003; 2003US-0463708P.  
XX 18-APR-2003; 2003US-0463716P.  
XX 18-APR-2003; 2003US-0463732P.  
XX 02-MAY-2003; 2003US-0467199P.  
XX 02-MAY-2003; 2003US-0467201P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 02-MAY-2003; 2003US-0467230P.

```
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RP, Huang MM, Kochakora S, Haishan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2804; 428bp; English.
XX
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX
XX Sequence 183 AA;
SQ
XX
XX Query Match 56.2%; Score 18; DB 8; Length 183;
XX Best Local Similarity 20.0%; Pred. No. 3.6;
XX Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 13
DB 36 CATTTTTC 45
XX
XX
XX RESULT 13
XX ABO76619
XX ID ABO76619 standard; protein, 191 AA.
XX
XX ABO76619;
AC
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #8794.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US651795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
```

```
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX
XX N-PSDB; ABD10190.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 25365; 455bp; English.
XX
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX
XX
XX Sequence 191 AA;
SQ
XX
XX Query Match 56.2%; Score 18; DB 7; Length 191;
XX Best Local Similarity 20.0%; Pred. No. 3.6;
XX Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXC 13
DB 49 CSTAASATSC 58
XX
XX
XX RESULT 14
XX ADZ56213
XX ID ADZ56213 standard; protein, 214 AA.
XX
XX ADZ56213;
AC
XX
XX 30-JUN-2005 (first entry)
XX
XX Human KIAA0779 splice variant clone CLN00149041.a, protein.
XX
XX KIAA0779; gene expression; inflammation; antiinflammatory; cancer;
XX proliferation; neoplasm; cytostatic; immune disorder; immunomodulator;
XX metabolic disorder; metabolic; viral infection; virucide; infection.
XX
XX Homo sapiens.
XX
XX MO2005035569-A2.
XX
XX 21-APR-2005.
XX
XX 12-OCT-2004; 2004WO-US033408.
XX
XX 10-OCT-2003; 2003US-0510612P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Wong JGP, Hestir K, Collins ALT;
XX
XX WPI; 2005-296268/30.
XX
```

DR N-PSDB; ADZ56200, ADZ56231.  
XX New isolated KIAA0779 nucleic acids and polypeptides, useful for  
PT diagnosing, preventing and/or treating inflammatory, immune, viral  
PT disorders and cancer, such as kidney, lung, ovarian, bladder, breast,  
PT prostate and skin cancers.  
XX  
XX Claim 11; SEQ ID NO 18; 121bp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule comprising at  
CC least one polynucleotide sequence (appearing as ADZ56196-ADZ56197,  
CC ADZ56199-ADZ56200, ADZ56226-ADZ56227 and ADZ56230-ADZ56231), sequences  
CC hybridizing to them under high stringency conditions, sequences having at  
CC least 80% sequence identity to them, their complements or biologically  
CC active fragments. The nucleic acids are splice variants of the human  
CC KIAA0779 gene, encoding transmembrane domain protein(s). Also included  
CC are a double-stranded isolated nucleic acid molecule comprising the  
CC nucleic acid molecule cited above, a vector comprising the isolated  
CC nucleic acid molecule cited above (and a promoter that regulates the  
CC expression of the nucleic acid molecule), a recombinant host cell  
CC comprising the nucleic acid molecule cited above, an isolated polypeptide  
CC (comprising the encoded proteins or the non-transmembrane (TM) regions of  
CC the proteins), a method of making a recombinant host cell, a method of  
CC making a polypeptide, a method of determining the presence of the nucleic  
CC acid molecule cited above in a sample, a method of determining the  
CC presence of a specific antibody to the polypeptide of (4) in a sample, a  
CC method of determining the presence of the polypeptide in a sample, an  
CC antibody specifically binding to and/or interfering with the biological  
CC activity of the nucleic acid molecule cited above (or the polypeptide or  
CC its biologically active fragment), a composition comprising a  
CC pharmaceutical carrier or excipient (and one or more active agents chosen  
CC from the nucleic acid molecule cited above, the vector, the polypeptide,  
CC and the antibody), a bacteriophage comprising the antibody (or its  
CC fragment), a bacterial cell comprising the bacteriophage, a recombinant  
CC host cell that produces the antibody, an animal injected with one or more  
CC active agents (chosen from the nucleic acid molecule, the vector, the  
CC host cell, the polypeptide, and the antibody), a diagnostic kit  
CC (comprising a nucleic acid molecule having at least 6 contiguous  
CC nucleotides from the nucleic acid molecule cited above, the isolated  
CC polypeptide, or antibody, and reagents to carry out an immunoassay), a  
CC method of making an antibody, a method of identifying a modulating agent  
CC that modulates the biological activity of the polypeptide, a modulator  
CC composition comprising a modulator and a pharmaceutical carrier, a method  
CC of treating a disease in a subject, and a method of treating cancer,  
CC preferably kidney, cervical, squamous lung, ovarian, bladder, breast,  
CC endometrial, prostate, and skin cancer, in a subject. The methods and  
CC compositions of the present invention are useful for the diagnosis,  
CC prevention and/or treatment of inflammatory, immune, viral disorders and  
CC cancer, such as kidney, cervical, squamous lung, ovarian, bladder,  
CC breast, endometrial, prostate, and skin cancer. The present sequence  
CC represents a protein encoded by a cDNA splice variant from the KIAA0779  
CC gene.  
XX  
XX Sequence 214 AA;  
SQ  
Query Match 56 2%; Score 18; DB 9; Length 214;  
Best Local Similarity 20.0%; Pred. No. 3.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXX 13  
DB 173 CAAAAAAC 182  
RESULT 15  
ADP30921  
ID ADP30921 standard; protein; 228 AA.  
XX  
XX ADP30921;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
DT  
XX  
DE Human secreted protein SEQ ID #1688.

XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
PD  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406589P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493441P.  
PR 08-AUG-2003; 2003US-0493470P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2919; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 228 AA;  
XX  
Query Match 56.2%; Score 18; DB 8; Length 228;  
Best Local Similarity 20.0%; Pred. No. 3.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXC 13  
DB 54 CATAAATTAC 63  
XX  
RESULT 16  
ADP31485  
ID ADP31485 standard; protein; 252 AA.  
XX  
AC ADP31485;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2252.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX PD 29-APR-2004.  
XX  
XX PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX

PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410963P.  
PR 17-SEP-2002; 2002US-0410964P.  
PR 17-SEP-2002; 2002US-0410965P.  
PR 17-SEP-2002; 2002US-0410966P.  
PR 17-SEP-2002; 2002US-0410967P.  
PR 17-SEP-2002; 2002US-0410968P.  
PR 17-SEP-2002; 2002US-0410969P.  
PR 17-SEP-2002; 2002US-0410970P.  
PR 17-SEP-2002; 2002US-0410971P.  
PR 17-SEP-2002; 2002US-0410972P.  
PR 17-SEP-2002; 2002US-0410973P.  
PR 17-SEP-2002; 2002US-0410974P.  
PR 17-SEP-2002; 2002US-0410975P.  
PR 17-SEP-2002; 2002US-0410976P.  
PR 17-SEP-2002; 2002US-0410977P.  
PR 17-SEP-2002; 2002US-0410978P.  
PR 17-SEP-2002; 2002US-0410979P.  
PR 17-SEP-2002; 2002US-0410980P.  
PR 17-SEP-2002; 2002US-0410981P.  
PR 17-SEP-2002; 2002US-0410982P.  
PR 17-SEP-2002; 2002US-0410983P.  
PR 17-SEP-2002; 2002US-0410984P.  
PR 17-SEP-2002; 2002US-0410985P.  
PR 17-SEP-2002; 2002US-0410986P.  
PR 17-SEP-2002; 2002US-0410987P.  
PR 17-SEP-2002; 2002US-0410988P.  
PR 17-SEP-2002; 2002US-0410989P.  
PR 17-SEP-2002; 2002US-0410990P.  
PR 17-SEP-2002; 2002US-0410991P.  
PR 17-SEP-2002; 2002US-0410992P.  
PR 17-SEP-2002; 2002US-0410993P.  
PR 17-SEP-2002; 2002US-0410994P.  
PR 17-SEP-2002; 2002US-0410995P.  
PR 17-SEP-2002; 2002US-0410996P.  
PR 17-SEP-2002; 2002US-0410997P.  
PR 17-SEP-2002; 2002US-0410998P.  
PR 17-SEP-2002; 2002US-0410999P.  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3483; 428bp; English.  
XX

CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 252 AA;  
Query Match 56.2%; Score 18; DB 8; Length 252;  
Best Local Similarity 20.0%; Pred. No. 3.9;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXX 13  
DB 210 CTTTATATAC 219  
RESULT 17  
ID ADP30479 standard; protein; 258 AA.  
XX  
AC ADP30479;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1246.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM Cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 09-JUN-2003; 2003US-047609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-048646P.  
PR 14-JUL-2003; 2003US-048646P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493373P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Halshan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 2477; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 258 AA;  
Query Match 56.2%; Score 18; DB 8; Length 258;  
Best Local Similarity 20.0%; Pred. No. 3.9;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXX 13  
DB 237 CTTTATATAC 246  
RESULT 18  
ADP31412

ID	ADP31412	standard; protein; 264 AA.
XX		
AC	ADP31412;	
XX		
DT	12-AUG-2004	(first entry)
XX		
DE	Human secreted protein SEQ ID #2179.	
XX		
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide	
KW	Cancer; Inflammatory; Immune;	human secreted protein.
XX		
OS	Homo sapiens.	
XX		
PN	WO2004035732-A2.	
XX		
PD	29-APR-2004.	
XX		
PF	28-AUG-2003;	2003WO-US026780.
XX		
PR	29-AUG-2002;	2002US-0406576P.
PR	29-AUG-2002;	2002US-0406579P.
PR	29-AUG-2002;	2002US-0406585P.
PR	29-AUG-2002;	2002US-0406588P.
PR	29-AUG-2002;	2002US-0406508P.
PR	29-AUG-2002;	2002US-0406611P.
PR	29-AUG-2002;	2002US-0406612P.
PR	29-AUG-2002;	2002US-0406616P.
PR	29-AUG-2002;	2002US-0406640P.
PR	29-AUG-2002;	2002US-0406642P.
PR	29-AUG-2002;	2002US-0406646P.
PR	29-AUG-2002;	2002US-0406653P.
PR	29-AUG-2002;	2002US-0406655P.
PR	29-AUG-2002;	2002US-0406666P.
PR	17-SEP-2002;	2002US-0410946P.
PR	17-SEP-2002;	2002US-0410947P.
PR	17-SEP-2002;	2002US-0410948P.
PR	17-SEP-2002;	2002US-0410949P.
PR	17-SEP-2002;	2002US-0410953P.
PR	17-SEP-2002;	2002US-0410957P.
PR	17-SEP-2002;	2002US-0410958P.
PR	17-SEP-2002;	2002US-0410959P.
PR	17-SEP-2002;	2002US-0410960P.
PR	17-SEP-2002;	2002US-0410961P.
PR	17-SEP-2002;	2002US-0410962P.
PR	17-SEP-2002;	2002US-0411019P.
PR	17-SEP-2002;	2002US-0411022P.
PR	17-SEP-2002;	2002US-0411023P.
PR	17-SEP-2002;	2002US-0411024P.
PR	17-SEP-2002;	2002US-0411032P.
PR	17-SEP-2002;	2002US-0411035P.
PR	17-SEP-2002;	2002US-0411037P.
PR	17-SEP-2002;	2002US-0411041P.
PR	17-SEP-2002;	2002US-0411045P.
PR	17-SEP-2002;	2002US-0411046P.
PR	17-SEP-2002;	2002US-0411048P.
PR	17-SEP-2002;	2002US-0411052P.
PR	17-SEP-2002;	2002US-0411055P.
PR	17-SEP-2002;	2002US-0411073P.
PR	17-SEP-2002;	2002US-0411082P.
PR	17-SEP-2002;	2002US-0411101P.
PR	17-SEP-2002;	2002US-0411104P.
PR	18-APR-2003;	2003US-0463708P.
PR	18-APR-2003;	2003US-0463708P.
PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US-0463732P.
PR	02-MAY-2003;	2003US-0467199P.
PR	02-MAY-2003;	2003US-0467201P.
PR	02-MAY-2003;	2003US-0467203P.
PR	02-MAY-2003;	2003US-0467230P.
PR	19-MAY-2003;	2003US-0471306P.
PR	19-MAY-2003;	2003US-0471366P.
PR	22-MAY-2003;	2003US-0472420P.
PR	22-MAY-2003;	2003US-0472430P.

PR	09-JUN-2003;	2003US-0476609P.
PR	09-JUN-2003;	2003US-0476641P.
PR	08-JUL-2003;	2003US-0465218P.
PR	08-JUL-2003;	2003US-0485223P.
PR	08-JUL-2003;	2003US-0485223P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486446P.
PR	14-JUL-2003;	2003US-0486480P.
PR	15-JUL-2003;	2003US-0486891P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.
XX		
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX		
PI	Williams LT, Chu K, Lee E, Heestir K, Beaurang PA, Behrens D,	
P1	Halenbeck RF, Huang WM, Kochakota S, Haishan L, Linemann T;	
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;	
XX		
XX	WPI; 2004-348436/32.	
DR		
XX		
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
PS	Claim 1; SEQ ID NO 3410; 428bp; English.	
CC		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
CC	available on WIPOWEB and is not in the specification.	
XX		
SQ	Sequence 264 AA;	
OY	Query Match	56.2%; Score 18; DB 8; Length 264;
	Best Local Similarity	20.0%; Pred. No. 3.9;
	Matches	2; Conservative
		0; Mismatches
		8; Indels
		0; Gaps
		0;
DB	4 CXXXXXXXXC 13	
	186 CATAAATTC 195	
RESULT 19		
ADP30500		
ID	ADP30500 standard; protein; 270 AA.	
XX		
AC	ADP30500;	
XX		
DT	12-AUG-2004 (first entry)	
XX		
DE	Human secreted protein SEQ ID #1267.	
XX		
KM	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;	
XX	Cancer; Inflammatory; Immune; human secreted protein.	
OS	Homo sapiens.	
XX		
PN	WO2004035732-A2.	
XX		
PD	29-APR-2004.	
XX		
PF	28-AUG-2003; 2003WO-US026780.	
XX		
PR	29-AUG-2002; 2002US-0406576P.	
PR	29-AUG-2002; 2002US-0406579P.	
PR	29-AUG-2002; 2002US-0406585P.	



PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486691P.  
PR 15-JUL-2003; 2003US-0486696P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T,  
XX Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;  
XX  
XX MPI; 2004-348438/32.

XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2498; 428pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 270 AA;  
XX  
Query Match . 56.2%; Score 18; DB 8; Length 270;  
Best Local Similarity 20.0%; Pred. No. 3.9;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
CY 4 CXXXXXXXXX 13  
DB 233 CTTATATATC 242  
XX  
RESULT 20  
ADP1321  
ID ADP1321 standard; protein; 270 AA.  
XX  
XX ADP1321;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #2088.  
DE  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.

17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467206P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D,  
P1 Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linnemann T;  
P1 Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI, 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3319; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
XX antinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present  
XX sequence represents a human secreted protein. The present sequence is  
XX available on WIPOMB and is not in the specification.  
XX  
XX Sequence 270 AA;

QY	13	XXXXXXXXXX
DB	254	CTTATATATAC
RESULT 21		
ID	ADP31473	ADP31473 standard; protein; 294 AA.
AC	ADP31473;	
XX		
DT	12-AUG-2004	(first entry)
DE	Human secreted protein seq	ID #2240.
XX		
KM	Cycostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide	
KW	Cancer; Inflammatory; Immune;	human secreted protein.
XX		
OS	Homo sapiens.	
XX		
PM	WO2004035732-A2.	
XX		
PD	29-APR-2004.	
XX		
PF	28-AUG-2003;	2003WO-US026780.
XX		
PR	29-AUG-2002;	2002US-0406576P.
PR	29-AUG-2002;	2002US-0406579P.
PR	29-AUG-2002;	2002US-0406585P.
PR	29-AUG-2002;	2002US-0406588P.
PR	29-AUG-2002;	2002US-0406608P.
PR	29-AUG-2002;	2002US-0406611P.
PR	29-AUG-2002;	2002US-0406612P.
PR	29-AUG-2002;	2002US-0406616P.
PR	29-AUG-2002;	2002US-0406640P.
PR	29-AUG-2002;	2002US-0406642P.
PR	29-AUG-2002;	2002US-0406646P.
PR	29-AUG-2002;	2002US-0406653P.
PR	29-AUG-2002;	2002US-0406655P.
PR	29-AUG-2002;	2002US-0406656P.
PR	17-SEP-2002;	2002US-0410944P.
PR	17-SEP-2002;	2002US-0410947P.
PR	17-SEP-2002;	2002US-0410948P.
PR	17-SEP-2002;	2002US-0410949P.
PR	17-SEP-2002;	2002US-0410953P.
PR	17-SEP-2002;	2002US-0410957P.
PR	17-SEP-2002;	2002US-0410958P.
PR	17-SEP-2002;	2002US-0410959P.
PR	17-SEP-2002;	2002US-0410960P.
PR	17-SEP-2002;	2002US-0410961P.
PR	17-SEP-2002;	2002US-0410962P.
PR	17-SEP-2002;	2002US-0411019P.
PR	17-SEP-2002;	2002US-0411022P.
PR	17-SEP-2002;	2002US-0411023P.
PR	17-SEP-2002;	2002US-0411024P.
PR	17-SEP-2002;	2002US-0411032P.
PR	17-SEP-2002;	2002US-0411035P.
PR	17-SEP-2002;	2002US-0411037P.
PR	17-SEP-2002;	2002US-0411041P.
PR	17-SEP-2002;	2002US-0411045P.
PR	17-SEP-2002;	2002US-0411046P.
PR	17-SEP-2002;	2002US-0411048P.
PR	17-SEP-2002;	2002US-0411052P.
PR	17-SEP-2002;	2002US-0411055P.
PR	17-SEP-2002;	2002US-0411073P.
PR	17-SEP-2002;	2002US-0411082P.
PR	17-SEP-2002;	2002US-0411101P.
PR	17-SEP-2002;	2002US-0411111P.
PR	18-APR-2003;	2003US-0463708P.
PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US-0463732P.
PR	02-MAY-2003;	2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

## (FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

## Claim 1; SEQ ID NO 3471; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.

## SQ Sequence 294 AA;

Query Match 56.2%; Score 18; DB 8; Length 294;

Best Local Similarity 20.0%; Pred. No. 4;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 72 CAATMAATC 81

## RESULT 22

ID ADP31192 standard; protein; 297 AA.

XX ADP31192;

XX 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #1959.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KM cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX WO2004035732-A2.

XX

PD 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

XX 17-SEP-2002; 2002US-0411046P.

XX 17-SEP-2002; 2002US-0411048P.

XX 17-SEP-2002; 2002US-0411052P.

XX 17-SEP-2002; 2002US-0411055P.

XX 17-SEP-2002; 2002US-0411073P.

XX 17-SEP-2002; 2002US-0411082P.

XX 17-SEP-2002; 2002US-0411101P.

XX 17-SEP-2002; 2002US-0411111P.

XX 18-APR-2003; 2003US-0463700P.

XX 18-APR-2003; 2003US-0463708P.

XX 18-APR-2003; 2003US-0463716P.

XX 02-MAY-2003; 2003US-0463732P.

XX 02-MAY-2003; 2003US-0467199P.

XX 02-MAY-2003; 2003US-0467201P.

XX 02-MAY-2003; 2003US-0467203P.

XX 02-MAY-2003; 2003US-0467230P.

XX 19-MAY-2003; 2003US-0471306P.

XX 19-MAY-2003; 2003US-0471306P.

XX 22-MAY-2003; 2003US-0472430P.

XX 09-JUN-2003; 2003US-0476609P.

XX 09-JUN-2003; 2003US-0476641P.

XX 08-JUL-2003; 2003US-0485218P.

XX 08-JUL-2003; 2003US-0485223P.

XX 08-JUL-2003; 2003US-0485224P.

XX 08-JUL-2003; 2003US-0485325P.

XX 14-JUL-2003; 2003US-0486446P.

XX 14-JUL-2003; 2003US-0486480P.

XX 15-JUL-2003; 2003US-0486891P.

XX 15-JUL-2003; 2003US-0486960P.

XX 08-AUG-2003; 2003US-0493341P.

XX 08-AUG-2003; 2003US-0493370P.

XX 08-AUG-2003; 2003US-0493573P.

XX 08-AUG-2003; 2003US-0493577P.

XX

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX Williams LT, Chu K, Lee E, Heslir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 XX WPI, 2004-348438/32.  
 DR  
 XX  
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 PS Claim 1; SEQ ID NO 3190; 428bp; English.  
 XX  
 CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMB and is not in the specification.  
 SQ Sequence 297 AA;  
 Query Match 56.2%; Score 18; DB 8; Length 297;  
 Best Local Similarity 20.0%; Pred. No. 4;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 4 CXXXXXXXXXC 13  
 Db 166 CTTATTATTC 175  
 RESULT 23  
 AAM25285  
 ID AAM25285 standard; protein; 328 AA.  
 XX  
 AC AAM25285;  
 XX  
 DT 16-OCT-2001 (first entry)  
 XX  
 DE Human protein sequence SEQ ID NO:800.  
 XX  
 KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
 KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
 KW antibacterial; endocrine; cardiac; central nervous system; virucide;  
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
 KW antiaggregant; haemostatic; vulnery; antitumor; osteopathic; eczema;  
 KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;  
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
 KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
 KW cardiac dysfunction; rheumatology; cardiac anaphylaxis; autoimmunity;  
 KW genetic disease; haemopoietic disorder; platelet disorder; asthma;  
 KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
 KW neurological disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153455-A2.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 22-DEC-2000; 2000WO-US035017.  
 XX  
 PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 XX  
 XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;  
 PI WPI; 2001-457603/49.  
 DR N-PSDB; AAH99226.  
 XX  
 PT Isolated human polynucleotides encoding polypeptides, useful for the  
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.  
 PS Claim 20; Page 187; 1217bp; English.  
 XX  
 CC AAH9166 to AAH99904 encode the human proteins given in AAM25225 to  
 CC AAM25963. The proteins can have activities based on the tissues and cells  
 CC they are expressed in, such as: antiinflammatory; antirheumatic;  
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;  
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
 CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;  
 CC antitumor; osteopathic; dermatological; antiallergic; antiasthmatic;  
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
 CC encoding them can be used in gene therapy, antisense therapy and vaccine  
 CC production. The proteins and polynucleotides are useful for screening for  
 CC agonists or antagonists of a protein and for the treatment and diagnosis  
 CC of disorders associated with the activity of a protein e.g. inflammation,  
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
 CC neurological disorders  
 SQ Sequence 328 AA;  
 Query Match 56.2%; Score 18; DB 4; Length 328;  
 Best Local Similarity 20.0%; Pred. No. 4.1;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 4 CXXXXXXXXXC 13  
 Db 63 CAAAAAAAC 72  
 RESULT 24  
 ADP30702  
 ID ADP30702 standard; protein; 339 AA.  
 XX  
 AC ADP30702;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #1469.  
 XX  
 KW Cytostatic; Antiinflammatory; immunosuppressive; Antibacterial; Virucide;  
 KW cancer; inflammatory; immune; human secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.

PT	29-AUG-2002;	2002US-0406642P.
PR	29-AUG-2002;	2002US-0406642P.
PR	29-AUG-2002;	2002US-0406646P.
PR	29-AUG-2002;	2002US-0406653P.
PR	29-AUG-2002;	2002US-0406653P.
PR	29-AUG-2002;	2002US-0406666P.
PR	17-SEP-2002;	2002US-0410946P.
PR	17-SEP-2002;	2002US-0410947P.
PR	17-SEP-2002;	2002US-0410948P.
PR	17-SEP-2002;	2002US-0410949P.
PR	17-SEP-2002;	2002US-0410953P.
PR	17-SEP-2002;	2002US-0410957P.
PR	17-SEP-2002;	2002US-0410958P.
PR	17-SEP-2002;	2002US-0410959P.
PR	17-SEP-2002;	2002US-0410960P.
PR	17-SEP-2002;	2002US-0410962P.
PR	17-SEP-2002;	2002US-0410962P.
PR	17-SEP-2002;	2002US-0411019P.
PR	17-SEP-2002;	2002US-0411022P.
PR	17-SEP-2002;	2002US-0411023P.
PR	17-SEP-2002;	2002US-0411024P.
PR	17-SEP-2002;	2002US-0411032P.
PR	17-SEP-2002;	2002US-0411035P.
PR	17-SEP-2002;	2002US-0411037P.
PR	17-SEP-2002;	2002US-0411041P.
PR	17-SEP-2002;	2002US-0411045P.
PR	17-SEP-2002;	2002US-0411046P.
PR	17-SEP-2002;	2002US-0411048P.
PR	17-SEP-2002;	2002US-0411052P.
PR	17-SEP-2002;	2002US-0411055P.
PR	17-SEP-2002;	2002US-0411073P.
PR	17-SEP-2002;	2002US-0411082P.
PR	17-SEP-2002;	2002US-0411101P.
PR	17-SEP-2002;	2002US-0411111P.
PR	18-APR-2003;	2003US-0463700P.
PR	18-APR-2003;	2003US-0463708P.
PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US-0463732P.
PR	02-MAY-2003;	2003US-0467199P.
PR	02-MAY-2003;	2003US-0467201P.
PR	02-MAY-2003;	2003US-0467203P.
PR	02-MAY-2003;	2003US-0467230P.
PR	19-MAY-2003;	2003US-0471306P.
PR	19-MAY-2003;	2003US-047136P.
PR	22-MAY-2003;	2003US-0472420P.
PR	22-MAY-2003;	2003US-0472430P.
PR	09-JUN-2003;	2003US-0476609P.
PR	09-JUN-2003;	2003US-0476641P.
PR	08-JUL-2003;	2003US-0485218P.
PR	08-JUL-2003;	2003US-0485223P.
PR	08-JUL-2003;	2003US-0485224P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486446P.
PR	14-JUL-2003;	2003US-0486480P.
PR	15-JUL-2003;	2003US-0486891P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.
XX		
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX		
PI	Williams LT,	Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI	Halebeck RF,	Huang MM, Gochakota S, Haibhan L, Linemann T;
PI	Pierce K,	Wang Y, Wong JGP, Wu G, Zhang H;
XX		
XX	WPI; 2004-348438/32.	
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
XX	genetic, bacterial and viral diseases.	
XX		

PS	Claim 1, SEQ ID NO 2700; 428bp; English.
CC	The present invention relates to an isolated nucleic acid molecule
XX	encoding a polypeptide which is believed to be cytostatic,
CC	antitumorigenic, immunosuppressive, antibacterial and virucidal. The
CC	composition and methods are useful for diagnosing, preventing and
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,
CC	immune, metabolic, genetic, bacterial and viral diseases. The present
CC	sequence represents a human secreted protein. The present sequence is
CC	available on WIPOMB and is not in the specification.
XX	
SQ	Sequence 339 AA;
OY	Query Match 56.2%; Score 18; DB 8; Length 339; Best Local Similarity 20.0%; Pred. No. 4.1; Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0
DB	4 CXXXXXXXXX 13 89 CAATTAATTTC 98
RESULT 25	
ID	ADP31441 standard; protein; 348 AA.
AC	ADP31441;
DT	12-AUG-2004 (first entry)
DE	Human secreted protein SEQ ID #2208.
XX	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW	cancer; Inflammatory; Immune; human secreted protein.
XX	Homo sapiens.
XX	WO2004035732-A2.
PN	29-APR-2004.
PD	28-AUG-2003; 2003WO-US026780.
PF	29-AUG-2002; 2002US-0406576P.
XX	29-AUG-2002; 2002US-0406579P.
PR	29-AUG-2002; 2002US-0406585P.
PR	29-AUG-2002; 2002US-0406588P.
PR	29-AUG-2002; 2002US-0406608P.
PR	29-AUG-2002; 2002US-0406611P.
PR	29-AUG-2002; 2002US-0406612P.
PR	29-AUG-2002; 2002US-0406616P.
PR	29-AUG-2002; 2002US-0406640P.
PR	29-AUG-2002; 2002US-0406642P.
PR	29-AUG-2002; 2002US-0406646P.
PR	29-AUG-2002; 2002US-0406653P.
PR	29-AUG-2002; 2002US-0406655P.
PR	29-AUG-2002; 2002US-0406666P.
PR	17-SEP-2002; 2002US-0410946P.
PR	17-SEP-2002; 2002US-0410947P.
PR	17-SEP-2002; 2002US-0410948P.
PR	17-SEP-2002; 2002US-0410949P.
PR	17-SEP-2002; 2002US-0410953P.
PR	17-SEP-2002; 2002US-0410957P.
PR	17-SEP-2002; 2002US-0410958P.
PR	17-SEP-2002; 2002US-0410959P.
PR	17-SEP-2002; 2002US-0410960P.
PR	17-SEP-2002; 2002US-0410961P.
PR	17-SEP-2002; 2002US-0410962P.
PR	17-SEP-2002; 2002US-0411019P.
PR	17-SEP-2002; 2002US-0411022P.
PR	17-SEP-2002; 2002US-0411023P.
PR	17-SEP-2002; 2002US-0411024P.
PR	17-SEP-2002; 2002US-0411032P.

CC	PR	17-SEP-2002;	2002US-0411035P.
CC	PR	17-SEP-2002;	2002US-0411037P.
CC	PR	17-SEP-2002;	2002US-0411041P.
CC	PR	17-SEP-2002;	2002US-0411045P.
CC	PR	17-SEP-2002;	2002US-0411046P.
CC	PR	17-SEP-2002;	2002US-0411052P.
CC	PR	17-SEP-2002;	2002US-0411058P.
CC	PR	17-SEP-2002;	2002US-0411055P.
CC	PR	17-SEP-2002;	2002US-0411073P.
CC	PR	17-SEP-2002;	2002US-0411082P.
CC	PR	17-SEP-2002;	2002US-0411011P.
CC	PR	18-APR-2003;	2003US-0463700P.
CC	PR	18-APR-2003;	2003US-0463708P.
CC	PR	18-APR-2003;	2003US-0463716P.
CC	PR	18-APR-2003;	2003US-0463732P.
CC	PR	02-MAY-2003;	2003US-0467199P.
CC	PR	02-MAY-2003;	2003US-0467201P.
CC	PR	02-MAY-2003;	2003US-0467203P.
CC	PR	02-MAY-2003;	2003US-0467230P.
CC	PR	19-MAY-2003;	2003US-0471336P.
CC	PR	19-MAY-2003;	2003US-0471336P.
CC	PR	22-MAY-2003;	2003US-0472420P.
CC	PR	22-MAY-2003;	2003US-0472430P.
CC	PR	09-JUN-2003;	2003US-0476609P.
CC	PR	09-JUN-2003;	2003US-0476641P.
CC	PR	08-JUL-2003;	2003US-0485218P.
CC	PR	08-JUL-2003;	2003US-0485223P.
CC	PR	08-JUL-2003;	2003US-0485224P.
CC	PR	08-JUL-2003;	2003US-0485252P.
CC	PR	14-JUL-2003;	2003US-0486446P.
CC	PR	14-JUL-2003;	2003US-0486480P.
CC	PR	15-JUL-2003;	2003US-0486891P.
CC	PR	15-JUL-2003;	2003US-0486960P.
CC	PR	08-AUG-2003;	2003US-0493341P.
CC	PR	08-AUG-2003;	2003US-0493370P.
CC	PR	08-AUG-2003;	2003US-0493573P.
CC	PR	08-AUG-2003;	2003US-0493577P.
CC	PR	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
CC	PA		
CC	XX		
CC	PI	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;	
CC	PI	Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;	
CC	PI	Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;	
CC	XX		
CC	DR	WPI; 2004-348438/32.	
CC	XX		
CC	PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
CC	PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
CC	PT	genetic, bacterial and viral diseases.	
CC	XX		
CC	PS	Claim 1; SEQ ID NO 3439; 428bp; English.	
CC	XX		
CC	CC	The present invention relates to an isolated nucleic acid molecule	
CC	CC	encoding a polypeptide which is believed to be cytostatic,	
CC	CC	anti-inflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	CC	composition and methods are useful for diagnosing, preventing and	
CC	CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	CC	sequence represents a human secreted protein. The present sequence is	
CC	CC	available on WIPOMB and is not in the specification.	

RESULT 26  
 ABG21039  
 ID ABG21039 standard; protein; 356 AA.  
 XX  
 AC ABG21039;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #21030.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HXSE-) HXSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS85226.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 51398; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridization probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 356 AA;

ID ADP31267 standard; protein; 357 AA.  
 AC ADP31267;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #2034.  
 XX  
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KM cancer; inflammatory; immune; human secreted protein.  
 OS Homo sapiens.  
 XX  
 PN MO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406642P.  
 PR 29-AUG-2002; 2002US-0406646P.  
 PR 29-AUG-2002; 2002US-0406653P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 29-AUG-2002; 2002US-0406666P.  
 PR 17-SRP-2002; 2002US-0410946P.  
 PR 17-SRP-2002; 2002US-0410947P.  
 PR 17-SRP-2002; 2002US-0410948P.  
 PR 17-SRP-2002; 2002US-0410949P.  
 PR 17-SRP-2002; 2002US-0410953P.  
 PR 17-SRP-2002; 2002US-0410957P.  
 PR 17-SRP-2002; 2002US-0410958P.  
 PR 17-SRP-2002; 2002US-0410959P.  
 PR 17-SRP-2002; 2002US-0410960P.  
 PR 17-SRP-2002; 2002US-0410961P.  
 PR 17-SRP-2002; 2002US-0410962P.  
 PR 17-SRP-2002; 2002US-0411019P.  
 PR 17-SRP-2002; 2002US-0411022P.  
 PR 17-SRP-2002; 2002US-0411023P.  
 PR 17-SRP-2002; 2002US-0411024P.  
 PR 17-SRP-2002; 2002US-0411032P.  
 PR 17-SRP-2002; 2002US-0411035P.  
 PR 17-SRP-2002; 2002US-0411037P.  
 PR 17-SRP-2002; 2002US-0411041P.  
 PR 17-SRP-2002; 2002US-0411045P.  
 PR 17-SRP-2002; 2002US-0411046P.  
 PR 17-SRP-2002; 2002US-0411048P.  
 PR 17-SRP-2002; 2002US-0411052P.  
 PR 17-SRP-2002; 2002US-0411055P.  
 PR 17-SRP-2002; 2002US-0411073P.  
 PR 17-SRP-2002; 2002US-0411082P.  
 PR 17-SRP-2002; 2002US-0411101P.  
 PR 17-SRP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-047136P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485255P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 XX  
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX  
 PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 XX  
 DR WPI; 2004-348438/32.  
 XX  
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX  
 PS Claim 1; SEQ ID NO 3265; 428bp; English.  
 XX  
 CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMB and is not in the specification.  
 XX  
 SQ Sequence 357 AA;  
 QY 4 CXXXXXXXXC 13  
 Db 14 CATATATATAC 23  
 RESULT 28  
 ID ADP30505 standard; protein; 357 AA.  
 XX  
 AC ADP30505;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #1272.  
 XX  
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KM cancer; inflammatory; immune; human secreted protein.  
 OS Homo sapiens.  
 XX  
 PN MO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI, 2004-348438/32.

XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2503; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 357 AA;  
Query Match 56.2%; Score 18; DB 8; Length 357;  
Best Local Similarity 20.0%; Pred. No. 4.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
OY 4 CXXXXXXXXXC 13  
DB 169 CTAATTTC 178  
RESULT 29  
ADP31439  
ID ADP31439 standard; protein; 360 AA.  
XX  
AC ADP31439;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2206.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.



PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476509P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 3437; 428bp; English.  
XX  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX CC encoding a polypeptide which is believed to be cytostatic,  
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX CC composition and methods are useful for diagnosing, preventing and  
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present  
XX CC sequence represents a human secreted protein. The present sequence is  
XX CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 360 AA;

Query Match 56.2%; Score 18; DB 8; Length 360;  
Best Local Similarity 20.0%; Pred. No. 4.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
Db 70 CTAATTAAC 79  
RESULT 30  
ADP31218  
ID ADP31218 standard; protein; 390 AA.  
XX  
XX ADP31218;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
XX  
XX  
XX Human secreted protein SEQ ID #1985.  
DE  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; Inflammatory; immune; human secreted protein.  
OS  
XX Homo sapiens.  
XX  
XX MO2004035732-A2.  
XX  
XX  
XX 29-APR-2004.  
XX  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411101P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 18-APR-2003; 2003US-0463700P.  
XX 18-APR-2003; 2003US-0463708P.  
XX 18-APR-2003; 2003US-0463716P.  
XX 18-APR-2003; 2003US-0463732P.  
XX 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485225P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486919P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1: SEQ ID NO 3216; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic, The  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 390 AA;  
XX  
Query Match 56.2%; Score 18; DB 8; Length 390;  
Best Local Similarity 20.0%; Pred. No. 4.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXC 13  
DB 60 CAAAAAATTC 69  
XX  
RESULT 31  
ADP31159  
ID ADP31159 standard; protein; 421 AA.  
XX  
AC ADP31159;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1926.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; Inflammatory; Immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX MO2004035732-AA2.  
XX  
XX

PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411038P.  
XX 17-SEP-2002; 2002US-0411039P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411101P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 18-APR-2003; 2003US-0463700P.  
XX 18-APR-2003; 2003US-0463708P.  
XX 18-APR-2003; 2003US-0463716P.  
XX 18-APR-2003; 2003US-0463732P.  
XX 02-MAY-2003; 2003US-0467199P.  
XX 02-MAY-2003; 2003US-0467201P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 19-MAY-2003; 2003US-0471306P.  
XX 19-MAY-2003; 2003US-0471336P.  
XX 22-MAY-2003; 2003US-0472420P.  
XX 22-MAY-2003; 2003US-0472430P.  
XX 09-JUN-2003; 2003US-0476609P.  
XX 09-JUN-2003; 2003US-0476641P.  
XX 08-JUL-2003; 2003US-0485218P.  
XX 08-JUL-2003; 2003US-0485223P.  
XX 08-JUL-2003; 2003US-0485224P.  
XX 08-JUL-2003; 2003US-0485225P.  
XX 14-JUL-2003; 2003US-0486446P.  
XX 14-JUL-2003; 2003US-0486480P.  
XX 15-JUL-2003; 2003US-0486919P.  
XX 15-JUL-2003; 2003US-0486960P.  
XX 08-AUG-2003; 2003US-0493341P.  
XX 08-AUG-2003; 2003US-0493370P.  
XX 08-AUG-2003; 2003US-0493573P.  
XX 08-AUG-2003; 2003US-0493577P.  
XX



XX Sequence 426 AA;  
SQ

Query Match 56.2%; Score 18; DB 8; Length 426;  
Best Local Similarity 20.0%; Pred. No. 4.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 66 CATTAAATTC 75

## RESULT 33

ABO82561  
ID ABO82561 standard; protein; 438 AA.

AC ABO82561;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #14736.

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI: 2003-615309/58.

XX N-PSDB; ABD16132.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 31307; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX

SQ Sequence 438 AA;

Query Match 56.2%; Score 18; DB 7; Length 438;  
Best Local Similarity 20.0%; Pred. No. 4.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 56 CSAATASSTC 65

## RESULT 34

IDP31465  
ID ADP31465 standard; protein; 453 AA.

AC ADP31465;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2232.

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX WO2004035732-A2.

XX 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411025P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411033P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

XX 17-SEP-2002; 2002US-0411046P.

XX 17-SEP-2002; 2002US-0411048P.

XX 17-SEP-2002; 2002US-0411052P.

XX 17-SEP-2002; 2002US-0411055P.

XX 17-SEP-2002; 2002US-0411073P.

XX 17-SEP-2002; 2002US-0411082P.

XX 17-SEP-2002; 2002US-0411101P.

XX 17-SEP-2002; 2002US-0411111P.

XX 18-APR-2003; 2003US-0463700P.

XX 18-APR-2003; 2003US-0463708P.

XX 18-APR-2003; 2003US-0463716P.

XX 18-APR-2003; 2003US-0463732P.

XX 02-MAY-2003; 2003US-0467199P.

XX 02-MAY-2003; 2003US-0467201P.

XX 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 05-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang WM, Kochakota S, Haislan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,  
XX  
DR WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 3463; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 453 AA;  
Query Match 56.2%; Score 18; DB 8; Length 453;  
Best Local Similarity 20.0%; Pred. No. 4.4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXX 13  
Db 312 CAAATTAATC 321  
RESULT 35  
ADP30854  
ID ADP30854 standard; protein; 471 AA.  
XX  
XX ADP30854;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
DT  
XX  
XX Human secreted protein SEQ ID #1621.  
DE  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
PD  
XX

PF 28-AUG-2003; 2003MO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 02-MAY-2003; 2003US-0467306P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX



Query Match 56.2%; Score 18; DB 8; Length 525;  
 Best Local Similarity 20.0%; Pred. No. 4.5;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 Db 475 CAAATTTAC 484

RESULT 37  
 ADP30855  
 ID ADP30855 standard; protein; 549 AA.  
 XX  
 AC ADP30855;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #1622.  
 XX  
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KM cancer; inflammatory; immune; human secreted protein.  
 OS Homo sapiens.  
 XX  
 PN MO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406642P.  
 PR 29-AUG-2002; 2002US-0406646P.  
 PR 29-AUG-2002; 2002US-0406653P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 29-AUG-2002; 2002US-0406666P.  
 PR 17-SEP-2002; 2002US-0410946P.  
 PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410948P.  
 PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410958P.  
 PR 17-SEP-2002; 2002US-0410959P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410961P.  
 PR 17-SEP-2002; 2002US-0410962P.  
 PR 17-SEP-2002; 2002US-0411019P.  
 PR 17-SEP-2002; 2002US-0411022P.  
 PR 17-SEP-2002; 2002US-0411023P.  
 PR 17-SEP-2002; 2002US-0411024P.  
 PR 17-SEP-2002; 2002US-0411032P.  
 PR 17-SEP-2002; 2002US-0411035P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411041P.  
 PR 17-SEP-2002; 2002US-0411045P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411101P.  
 PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471336P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485244P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493541P.  
 PR 08-AUG-2003; 2003US-0493570P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Lahnemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 XX WPI, 2004-348438/32.  
 DR  
 XX  
 XX  
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 PS Claim 1; SEQ ID NO 2853; 428bp; English.  
 XX  
 XX The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOWEB and is not in the specification.  
 XX  
 SO Sequence 549 AA;

Query Match 56.2%; Score 18; DB 8; Length 549;  
 Best Local Similarity 20.0%; Pred. No. 4.6;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 Db 220 CAAATATATC 229

RESULT 38  
 ADP31416  
 ID ADP31416 standard; protein; 555 AA.  
 XX  
 AC ADP31416;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #2183.  
 XX  
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KM cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.  
OS  
XX  
XX WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410955P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410969P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476611P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485244P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PA  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGE, Wu G, Zhang H;  
XX MPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3414; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 555 AA;  
XX  
Query Match 56.2%; Score 18; DB 8; Length 555;  
Best Local Similarity 20.0%; Pred. No. 4.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXX 13  
DB 211 CTTTATTTC 220  
RESULT 39  
ID ADP31417 standard; protein; 555 AA.  
ADP31417  
AC ADP31417;  
XX  
XX 12-AUG-2004 (first entry)  
DT XX  
XX  
DE Human secreted protein SEQ ID #2184.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.



PR 29-AUG-2002; 2002US-040666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-048525P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-049341P.  
PR 08-AUG-2003; 2003US-049370P.  
PR 08-AUG-2003; 2003US-049373P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beauregard PA, Behrens D,  
PI Halsebeck RF, Huang MM, Kochakota S, Haislan L, Lannemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. Cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3415; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
XX SQ Sequence 555 AA;  
Query Match 56.2%; Score 18; DB 8; Length 555;  
Best Local Similarity 20.0%; Pred No. 4.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXC 13  
Db 211 CTTATTTC 220  
RESULT 40  
ADP31699  
ID ADP31699 standard; protein; 588 AA.  
XX  
AC ADP31699;  
XX  
XX 12-AUG-2004 (first entry)  
DT  
XX  
DE Human secreted protein SEQ ID #2466.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 29-AUG-2002; 2002US-040666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411010P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-047136P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-047609P.  
 PR 09-JUN-2003; 2003US-047641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.

## (FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Halenbeck RP, Huang MM, Kochakota S, Haislan L, Linnemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

DR WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.

PS Claim 1; SEQ ID NO 3697; 428bp; English.

CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMB and is not in the specification.

XX SQ Sequence 588 AA;

Query Match 56.2%; Score 18; DB 8; Length 588;  
 Best Local Similarity 20.0%; Pred. No. 4.6;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXC 13  
 Db 277 CATTTC 286

RESULT 41  
 AD56209  
 ID AD56209 standard; protein; 591 AA.  
 XX  
 AC AD56209;

XX 30-JUN-2005 (first entry)  
 DT Human KIAA0779 splice variant clone 24980850:24980849a, nonTM\_1 #1.  
 XX  
 DE KIAA0779; gene expression; inflammation; antiinflammatory; cancer;  
 XX proliferation; neoplasm; cytostatic; immune disorder; immunomodulator;  
 KW metabolic disorder; metabolic; viral infection; virucide; infection.  
 KW  
 XX Homo sapiens.  
 OS  
 PN MO2005035569-A2.  
 XX  
 PD 21-APR-2005.  
 XX  
 PF 12-OCT-2004; 2004WO-US033408.  
 XX  
 PR 10-OCT-2003; 2003US-0510612P.  
 XX  
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX  
 PI Wong JGP, Hestir K, Collins ALT;  
 XX  
 DR WPI; 2005-296268/30.  
 XX  
 PT New isolated KIAA0779 nucleic acids and polypeptides, useful for  
 PT diagnosing, preventing and/or treating inflammatory, immune, viral  
 PT disorders and cancer, such as kidney, lung, ovarian, bladder, breast,  
 PT prostate and skin cancers.  
 XX  
 PS Claim 11; SEQ ID NO 14; 121bp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule comprises at  
 CC least one polynucleotide sequence (appearing as AD56196-AD56197,  
 CC AD56199-AD56200, AD56226-AD56227 and AD56230-AD56231), sequences  
 CC hybridizing to them under high stringency conditions, sequences having at  
 CC least 80% sequence identity to them, their complements or biologically  
 CC active fragments. The nucleic acids are splice variants of the human  
 CC KIAA0779 gene, encoding transmembrane domain protein(s). Also included  
 CC are a double-stranded isolated nucleic acid molecule comprising the  
 CC nucleic acid molecule cited above, a vector comprising the isolated  
 CC nucleic acid molecule cited above, and a promoter that regulates the  
 CC expression of the nucleic acid molecule(s), a recombinant host cell  
 CC comprising the nucleic acid molecule cited above, an isolated polypeptide  
 CC (comprising the encoded proteins or the non-transmembrane (TM) regions of  
 CC the proteins), a method of making a recombinant host cell, a method of  
 CC making a polypeptide, a method of determining the presence of the nucleic  
 CC acid molecule cited above in a sample, a method of determining the  
 CC presence of a specific antibody to the polypeptide of (4) in a sample, a  
 CC method of determining the presence of the polypeptide in a sample, an  
 CC antibody specifically binding to and/or interfering with the biological  
 CC activity of the nucleic acid molecule cited above (or the polypeptide or  
 CC its biologically active fragment), a composition comprising a  
 CC pharmaceutical carrier or excipient (and one or more active agents chosen  
 CC from the nucleic acid molecule cited above, the vector, the polypeptide,  
 CC and the antibody), a bacteriophage comprising the antibody (or its  
 CC fragment), a bacterial cell comprising the bacteriophage, a recombinant  
 CC host cell that produces the antibody, an animal injected with one or more  
 CC active agents (chosen from the nucleic acid molecule, the vector, the  
 CC host cell, the polypeptide, and the antibody), a diagnostic kit  
 CC (comprising a nucleic acid molecule having at least 6 contiguous  
 CC nucleotides from the nucleic acid molecule cited above, the isolated  
 CC polypeptide, or antibody, and reagents to carry out an immunoassay), a  
 CC method of making an antibody, a method of identifying a modulating agent  
 CC that modulates the biological activity of the polypeptide, a modulator  
 CC composition comprising a modulator and a pharmaceutical carrier, a method  
 CC of treating a disease in a subject, and a method of treating cancer,  
 CC preferably kidney, cervical, squamous lung, ovarian, bladder, breast,  
 CC endometrial, prostate, and skin cancer, in a subject. The methods and  
 CC compositions of the present invention are useful for the diagnosis,  
 CC prevention and/or treatment of inflammatory, immune, viral disorders and  
 CC cancer, such as kidney, cervical, squamous lung, ovarian, bladder,  
 CC breast, endometrial, prostate, and skin cancer. The present sequence

CC represents a non TM region of a protein encoded by a cDNA splice variant  
 CC from the KIAA0779 gene.  
 XX  
 XX  
 SQ Sequence 591 AA;  
 Query Match 56.2%; Score 18; DB 9; Length 591;  
 Best Local Similarity 20.0%; Pred. No. 4.6;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 4 CXXXXXXXXX 13  
 DB 173 CAAAAAAC 182  
 RESULT 42  
 AD256223  
 ID AD256223 standard; protein; 591 AA.  
 XX  
 AC AD256223;  
 XX  
 DT 30-JUN-2005 (first entry)  
 DE Human KIAA0779 splice variant clone 24980850:24980849a, nonTM\_1 #2.  
 XX  
 KM KIAA0779; gene expression; inflammation; antiinflammatory; cancer;  
 KM proliferation; neoplasm; cytostatic; immune disorder; immunomodulator;  
 KM metabolic disorder; metabolic; viral infection; virucide; infection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO2005035569-A2.  
 XX  
 PD 21-APR-2005.  
 PF 12-OCT-2004; 2004MO-US033408.  
 XX  
 PR 10-OCT-2003; 2003US-0510612P.  
 XX  
 PA (PIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX  
 PI Wong JGP, Hestir K, Collins ALT;  
 XX  
 DR WPI; 2005-296268/30.  
 XX  
 PT New isolated KIAA0779 nucleic acids and polypeptides, useful for  
 PT diagnosing, preventing and/or treating inflammatory, immune, viral  
 PT disorders and cancer, such as kidney, lung, ovarian, bladder, breast,  
 PT prostate and skin cancers.  
 XX  
 PS Claim 11, SEQ ID NO 28; 121bp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule comprises at  
 CC least one polynucleotide sequence (appearing as AD256196-AD256197,  
 CC AD256199-AD256200, AD256226-AD256227 and AD256230-AD256231), sequences  
 CC hybridizing to them under high stringency conditions, sequences having at  
 CC least 80% sequence identity to them, their complements or biologically  
 CC active fragments. The nucleic acids are splice variants of the human  
 CC KIAA0779 gene, encoding transmembrane domain protein(s). Also included  
 CC are a double-stranded isolated nucleic acid molecule comprising the  
 CC nucleic acid molecule cited above, a vector comprising the isolated  
 CC nucleic acid molecule cited above, (and a promoter that regulates the  
 CC expression of the nucleic acid molecule), a recombinant host cell  
 CC comprising the nucleic acid molecule cited above, an isolated polypeptide  
 CC (comprising the encoded proteins or the non-transmembrane (TM) regions of  
 CC the proteins), a method of making a recombinant host cell, a method of  
 CC making a polypeptide, a method of determining the presence of the nucleic  
 CC acid molecule cited above in a sample, a method of determining the  
 CC presence of a specific antibody to the polypeptide of (4) in a sample, a  
 CC method of determining the presence of the polypeptide in a sample, an  
 CC antibody specifically binding to and/or interfering with the biological  
 CC activity of the nucleic acid molecule cited above (or the polypeptide or  
 CC its biologically active fragment), a composition comprising a  
 CC pharmaceutical carrier or excipient (and one or more active agents chosen

CC from the nucleic acid molecule cited above, the vector, the polypeptide,  
 CC and the antibody), a bacteriophage comprising the antibody (or its  
 CC fragment), a bacterial cell comprising the bacteriophage, a recombinant  
 CC host cell that produces the antibody, an animal injected with one or more  
 CC active agents (chosen from the nucleic acid molecule, the vector, the  
 CC host cell, the polypeptide, and the antibody), a diagnostic kit  
 CC (comprising a nucleic acid molecule having at least 6 contiguous  
 CC nucleotides from the nucleic acid molecule cited above, the isolated  
 CC polypeptide, or antibody, and reagents to carry out an immunoassay), a  
 CC method of making an antibody, a method of identifying a modulating agent  
 CC that modulates the biological activity of the polypeptide, a modulator  
 CC composition comprising a modulator and a pharmaceutical carrier, a method  
 CC of treating a disease in a subject, and a method of treating cancer,  
 CC preferably kidney, cervical, squamous lung, ovarian, bladder, breast,  
 CC endometrial, prostate, and skin cancer, in a subject. The methods and  
 CC compositions of the present invention are useful for the diagnosis,  
 CC prevention and/or treatment of inflammatory, immune, viral disorders and  
 CC cancer, such as kidney, cervical, squamous lung, ovarian, bladder,  
 CC breast, endometrial, prostate, and skin cancer. The present sequence  
 CC represents a non TM region of a protein encoded by a cDNA splice variant  
 CC from the KIAA0779 gene.  
 XX  
 SQ Sequence 591 AA;  
 Query Match 56.2%; Score 18; DB 9; Length 591;  
 Best Local Similarity 20.0%; Pred. No. 4.6;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 4 CXXXXXXXXX 13  
 DB 173 CAAAAAAC 182  
 RESULT 43  
 ADP30865  
 ID ADP30865 standard; protein; 600 AA.  
 XX  
 AC ADP30865;  
 XX  
 DT 12-AUG-2004 (first entry)  
 DE Human secreted protein SEQ ID #1632.  
 XX  
 DB Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KM cancer; inflammatory; immune; human secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003MO-US026780.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406589P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406642P.  
 PR 29-AUG-2002; 2002US-0406646P.  
 PR 29-AUG-2002; 2002US-0406649P.  
 PR 29-AUG-2002; 2002US-0406653P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 29-AUG-2002; 2002US-0406666P.  
 PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410948P.  
 PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410953P.



PR 17-SEP-2002; 2002US-041111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471336P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.

## (FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 XX WPI; 2004-348438/32.

DR New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX

PS Claim 1, SEQ ID NO 2938; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPWEB and is not in the specification.  
 XX

SO Sequence 604 AA;

Query Match 56.2%; Score 18; DB 8; Length 604;

Best Local Similarity 20.0%; Pred. No. 4.7;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 Db 430 CTTATTAAAC 439

## RESULT 45

ADP30941  
 ID ADP30941 standard; protein; 604 AA.

ADP30941;

12-AUG-2004 (first entry)

Human secreted protein SEQ ID #1708.

Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX MO2004035732-A2.

PD 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

PR 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406640P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410961P.

PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.

PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.

PR 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411101P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485325P.

PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2939; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC anti-inflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 604 AA:  
Query Match 56.2%; Score 18; DB 8; Length 604;  
Best Local Similarity 20.0%; Pred. No. 4.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXX 13  
DB 430 CTTATTAAAC 439  
RESULT 46  
ADP30507  
ID ADP30507 standard; protein; 605 AA.  
XX  
XX ADP30507;  
AC  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1274.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
OS  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-048525P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2505; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,

CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.

XX  
SQ Sequence 605 AA;

Query Match 56.2%; Score 18; DB 8; Length 605;  
Best Local Similarity 20.0%; Pred. No. 4.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 570 CAAATTATTC 579

#### RESULT 47

ADP31657  
ID ADP31657 standard; protein; 617 AA.

XX ADP31657;

XX 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #2424.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

XX WO2004035732-A2.

XX 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411011P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463709P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 14-JUL-2003; 2003US-0486452P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486860P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LR, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;

PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Lunemann T;

PI Pierce K, Wang Y, Wong JCF, Wu G, Zhang H;

DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3655; 428pp; English.

XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic.  
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present  
XX sequence represents a human secreted protein. The present sequence is  
XX available on WIPOMEB and is not in the specification.

XX  
SQ Sequence 617 AA;

Query Match 56.2%; Score 18; DB 8; Length 617;  
Best Local Similarity 20.0%; Pred. No. 4.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 67 CAAATTATAC 76

#### RESULT 48

ABR41596  
ID ABR41596 standard; protein; 626 AA.

XX

AC ABR41596;  
XX  
XX 02-JUN-2003 (first entry)  
XX  
DE Human DITHP antigen recognition protein.  
XX  
XX  
XX Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;  
XX cancer; cell proliferative disorder; autoimmune disorder;  
XX inflammatory disorder; infection; hormonal disorder; metabolic disorder;  
XX neurological disorder; gastrointestinal disorder; transport disorder;  
XX connective tissue disorder; drug screening; proteome analysis;  
XX gene therapy; antisense therapy; genotyping; transgenic animal; knock in;  
XX disease model; toxicological testing; transcript imaging;  
XX antigen recognition.  
XX  
XX Homo sapiens.  
XX  
XX WO200297031-A2.  
XX  
XX 05-DEC-2002.  
XX  
XX 27-MAR-2002; 2002WO-US010056.  
XX  
XX 28-MAR-2001; 2001US-0279619P.  
XX 29-MAR-2001; 2001US-0280067P.  
XX 29-MAR-2001; 2001US-0280068P.  
XX 16-MAY-2001; 2001US-0291280P.  
XX 17-MAY-2001; 2001US-0291829P.  
XX 17-MAY-2001; 2001US-0291849P.  
XX 19-JUN-2001; 2001US-0299428P.  
XX 20-JUN-2001; 2001US-0299776P.  
XX 20-JUN-2001; 2001US-0300001P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
XX Dufour GE, Hillman JL, Yu JY, Thason O, Yap PE, Amshay SR;  
XX Daugherty SC, Dam TC, Nguyen DA, Kleefeld Y, Gerstin EH;  
XX Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
XX Flores V, Marwaha R, Lo A, Ian RV, Urashka ME;  
XX  
XX MPI: 2003-129518/12.  
XX N-PSDB; ACC46534.  
XX  
XX Novel human diagnostic and therapeutic polypeptide useful for identifying  
XX test compound which specifically binds to a polypeptide encoded by human  
XX diagnostic and therapeutic polynucleotide, and to induce antibodies.  
XX  
XX Claim 27; SEQ ID NO 1131; 591pp; English.  
XX  
XX The invention relates to novel human diagnostic and therapeutic  
XX polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded  
XX proteins (DITHP; ABR41136-ABR41812). The invention also relates to  
XX polynucleotide sequences at least 90% identical to the dithp cDNA  
XX sequences of the invention; recombinant vectors, host cells and  
XX transgenic organisms comprising a dithp nucleic acid sequence; the  
XX recombinant production of DITHP proteins; antibodies specific for DITHP  
XX proteins; microarrays comprising dithp nucleic acid sequences; methods of  
XX detecting dithp nucleotide and protein sequences; methods of screening  
XX for compounds which specifically bind a DITHP protein; and methods of  
XX assessing the toxicity of test compounds using a dithp hybridisation  
XX probe. Dithp nucleic acid sequences and DITHP proteins may be used in the  
XX diagnosis of a wide variety of conditions including cancer and other cell  
XX proliferative disorders; autoimmune or inflammatory disorders; bacterial,  
XX viral, fungal or parasitic infections; hormonal disorders; metabolic  
XX disorders; neurological disorders; gastrointestinal disorders; transport  
XX disorders; and connective tissue disorders. They may also be used to  
XX screen for modulators of protein activity or gene expression. DITHP  
XX proteins can additionally be used in analysis of the proteome of a tissue  
XX or cell type and to induce antibodies. The dithp nucleic acids are  
XX additionally useful in somatic or germline gene therapy of the disorders  
XX mentioned above, as a source of antisense sequences, as a source of  
XX probes and primers, in genotyping and identification of individuals, in

CC the generation of transgenic animal models of human disease or knock in  
CC humanised animals, in toxicological testing, and in transcript imaging.  
CC The present sequence represents a DITHP protein which has antigen  
CC recognition activity. Note: The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX SQ Sequence 626 AA;  
XX  
XX Query Match 56.2%; Score 18; DB 6; Length 626;  
XX Best Local Similarity 20.0%; Pred. No. 4.7;  
XX Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
XX  
XX QY 4 CXXXXXXXC 13  
XX  
XX Db 41 CTSTSSAASC 50  
XX  
XX RESULT 49  
XX ADP30858  
XX ID ADP30858 standard; protein; 645 AA.  
XX AC ADP30858;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
XX DE Human secreted protein SEQ ID #1625.  
XX  
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
XX OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX PD 29-APR-2004.  
XX  
XX PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.



PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411101P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467210P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471316P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476611P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 14-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486466P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.

# (FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Halendeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

DR MPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 2856; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytotoxic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMB and is not in the specification.

XX Sequence 645 AA;

Query Match 56.2%; Score 18; DB 8; Length 645;  
 Best Local Similarity 20.0%; Pred. NO. 4.7;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 4 CXXXXXXXXXC 13  
 DB 161 CAAATTTTTC 170

RESULT 50  
 AAY94907  
 ID AAY94907 standard; protein; 653 AA.

AC AAY94907;  
 XX 16-JUN-2000 (first entry)  
 DT Human secreted protein clone ca106\_19x protein sequence SEQ ID NO:20.  
 DE  
 XX Human; secreted protein; immunostimulant; immunosuppressant; virucide;  
 KW antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;  
 KW antidiabetic; antiaslathmatic; antiarthritic; antirheumatic; protozoacide;  
 KW antithyroid; immune deficiency; severe combined immunodeficiency; SCID;  
 KW infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;  
 KW connective tissue disease; multiple sclerosis; erythematosis;  
 KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;  
 KW Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;  
 KW insulin dependent diabetes mellitus; graft-versus-host-disease;  
 KW autoimmune inflammatory eye disease; allergy.  
 OS Homo sapiens.  
 OS  
 OS  
 PN WO200009552-A1.  
 PN  
 PD 24-FEB-2000.  
 XX  
 XX 13-AUG-1999; 99WO-US018298.  
 PF  
 XX 14-AUG-1998; 98US-0096622P.  
 PR 17-AUG-1998; 98US-0096815P.  
 PR 04-SEP-1998; 98US-0099229P.  
 PR 23-OCT-1998; 98US-0105368P.  
 PR 08-JAN-1999; 99US-0115234P.  
 PR 12-FEB-1999; 99US-0119931P.  
 PR 18-FEB-1999; 99US-0120575P.  
 PR 30-APR-1999; 99US-0132020P.  
 PR 11-AUG-1999; 99US-0148424P.  
 XX  
 XX (GEMV ) GENETICS INST INC.  
 PA  
 XX  
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Agostino MJ, Steinger RJ, Spaulding V;  
 PI Wong GG, Clark HF, Reichen K;  
 XX  
 XX MPI; 2000-205979/18.  
 DR  
 XX  
 XX New polynucleotides encoding secreted proteins, which may have e.g.  
 PT nutritional, chemokine, immune stimulating or suppressing, hematopoiesis  
 PT regulating, tissue growth, activin/inhibin antiinflammatory or tumor  
 PT inhibition activity.  
 PT  
 XX Claim 29; Page 487-489; 641pp; English.  
 XX  
 XX AAI1618 to AAI1697 encode the human secreted proteins given in AAY94898  
 CC to AAY94980, isolated from human adult brain, adult thyroid, adult  
 CC retina, foetal carcinoma, adult blood, adult neural, foetal kidney, adult  
 CC placenta, adult testis, whole embryo, adult cartilage, kidney, foetal  
 CC brain, adult thymus, foetal placenta, adult uterus, adult tumour, and  
 CC adult bladder, cDNA libraries. The polynucleotides and proteins are  
 CC predicted to have biological activities which would make them suitable  
 CC for treating, preventing or ameliorating medical conditions in humans and  
 CC animals. The polynucleotides can be used as markers for tissues in which  
 CC the protein is preferentially expressed, as molecular weight markers on  
 CC Southern gels, and as chromosome markers or tags to identify chromosomes  
 CC or to map gene positions. The proteins can be used in the treatment of  
 CC immune deficiencies and disorders, such as severe combined  
 CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other  
 CC infections. These infections include human immunodeficiency virus (HIV),  
 CC hepatitis, herpesviruses, mycobacteria, leishmania spp., malaria and  
 CC candidiasis. The proteins can be used to treat autoimmune disorders such  
 CC as connective tissue disease, multiple sclerosis, systemic lupus  
 CC erythematosis, rheumatoid arthritis, autoimmune pulmonary inflammation,  
 CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent  
 CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and  
 CC autoimmune inflammatory eye disease. The proteins can also be used to  
 CC treat allergic conditions, such as asthma. AAI1698 to AAI16774 represent

CC probes for the human secreted proteins from the present invention  
 XX  
 SQ Sequence 653 AA;

Query Match 56.2%; Score 18; DB 3; Length 653;  
 Best Local Similarity 20.0%; Pred. No. 4.7;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXC 13  
 |  
 Db 173 CAAAAAAC 182

Search completed: January 4, 2006, 15:56:03  
 Job time : 122.496 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 16.5565 Seconds  
(without alignments)  
92.983 Million cell updates/sec

Title: US-09-932-322-4  
Perfect score: 32  
Sequence: 1 XXXXXXXXXXXXXXXX 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 1000 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	56.2	313	2	SS9448
2	18	56.2	480	2	E75433
3	17	53.1	194	2	JQ1560
4	17	53.1	220	2	JC4082
5	17	53.1	376	2	AB3920
6	17	53.1	388	2	C97694
7	17	53.1	389	2	T23167
8	17	53.1	379	2	T37314
9	17	53.1	634	2	S33575
10	17	53.1	942	2	D87803
11	17	53.1	2704	2	S09118
12	17	50.0	71	2	F84312
13	16	50.0	85	2	E70531
14	16	50.0	104	2	S51479
15	16	50.0	105	2	A71249
16	16	50.0	108	2	F72549
17	16	50.0	126	2	A23473
18	16	50.0	129	2	T49498
19	16	50.0	133	2	A96746
20	16	50.0	147	2	S09762
21	16	50.0	153	2	PN0103
22	16	50.0	161	2	T28088
23	16	50.0	164	2	JQ1252
24	16	50.0	229	2	T34277
25	16	50.0	242	2	A45724
26	16	50.0	244	2	S72219
27	16	50.0	251	2	A55035
28	16	50.0	262	1	JQ1724
29	16	50.0	263	2	S47537

30	16	50.0	275	2	A85856	probable elongatio
31	16	50.0	275	2	B64986	hypothetical 30.9
32	16	50.0	275	2	G91011	probable elongatio
33	16	50.0	294	2	T23682	hypothetical prote
34	16	50.0	306	2	C71498	probable yop trans
35	16	50.0	317	2	T00500	probable elicitor
36	16	50.0	321	2	A84792	hypothetical prote
37	16	50.0	349	2	JR0202	paired-box-contain
38	16	50.0	380	2	T04007	hypothetical prote
39	16	50.0	397	2	T25077	hypothetical prote
40	16	50.0	422	2	S41514	hypothetical prote
41	16	50.0	473	2	C81039	RAD52 protein homo
42	16	50.0	473	2	C81984	lipopolysaccharide
43	16	50.0	475	2	T36342	probable lipopolys
44	16	50.0	486	2	B86460	probable glutamate
45	16	50.0	510	2	A55207	hypothetical prote
46	16	50.0	537	2	G31277	glycerol-3-phospha
47	16	50.0	580	2	B38418	guanine transport
48	16	50.0	638	2	T22518	jockey protein 1 -
49	16	50.0	642	2	C81345	hypothetical prote
50	16	50.0	648	2	T37581	hypothetical prote
51	16	50.0	708	1	TFMOL	probable serine-th
52	16	50.0	708	2	UC2323	lactotransferrin p
53	16	50.0	711	1	TFMUL	lactotransferrin p
54	16	50.0	907	2	A86460	99.9K hypothetical
55	16	50.0	975	2	T29308	hypothetical prote
56	16	50.0	1031	2	T06130	hypothetical prote
57	16	50.0	1069	2	T22138	hypothetical prote
58	16	50.0	1105	2	T22132	hypothetical prote
59	16	50.0	1136	2	AB1581	hypothetical prote
60	16	50.0	1380	2	T18309	receptor-adenylate
61	16	50.0	1474	2	D88550	protein ZC84.6 [lm
62	16	50.0	2233	2	T28669	surface protein 51
63	16	50.0	2395	2	S50820	surface protein ty
64	16	50.0	2533	2	T28675	alpha-51D immobili
65	16	50.0	2533	2	T28674	alpha-51D-immobil
66	16	50.0	2543	2	T31687	surface antigen - P
67	16	50.0	2718	2	A23475	G surface protein
68	15	46.9	40	1	SMRF	metallothionein Mt
69	15	46.9	40	2	B61194	metallothionein Mt
70	15	46.9	43	2	S33483	hypothetical prote
71	15	46.9	74	2	A82210	hypothetical prote
72	15	46.9	93	2	T18118	hypothetical prote
73	15	46.9	109	2	S64309	hypothetical prote
74	15	46.9	111	2	A72079	hypothetical prote
75	15	46.9	111	2	B86544	hypothetical prote
76	15	46.9	113	2	T10136	hypothetical prote
77	15	46.9	123	2	B81120	probable periplasm
78	15	46.9	124	2	C81707	hypothetical prote
79	15	46.9	131	2	T36866	probable iron-sulf
80	15	46.9	135	2	T38956	very hypothetical
81	15	46.9	139	2	D97348	hypothetical prote
82	15	46.9	141	2	A80029	Merr-family transc
83	15	46.9	141	2	PC1294	trophozoite surfac
84	15	46.9	146	2	T01683	immobilization ant
85	15	46.9	149	2	AF2977	conserved hypothet
86	15	46.9	158	2	T14975	hypothetical prote
87	15	46.9	160	2	F98305	hypothetical prote
88	15	46.9	169	1	S18946	ultra high-sulfur
89	15	46.9	171	2	T10315	hypothetical prote
90	15	46.9	179	2	T22471	hypothetical prote
91	15	46.9	181	2	T31481	hypothetical prote
92	15	46.9	185	2	T51003	hypothetical prote
93	15	46.9	189	2	B80969	probable tail asse
94	15	46.9	191	2	S69735	hypothetical prote
95	15	46.9	195	2	T19617	hypothetical prote
96	15	46.9	199	2	C90734	probable outer mem
97	15	46.9	199	2	B85584	hypothetical prote
98	15	46.9	203	2	H81695	recombination prot
99	15	46.9	204	2	A28864	proteinase B (EC 3
100	15	46.9	207	2	A91177	Shut-like protein
101	15	46.9	207	2	B86023	hypothetical prote
102	15	46.9	208	2	T07732	tuberculosis-relat

103	15	46.9	211	2	S34274	probable secretory
104	15	46.9	211	2	UC4883	scyllallopepsin B
105	15	46.9	212	2	T22463	hypothetical prote
106	15	46.9	214	2	C70812	probable 1pqq prote
107	15	46.9	219	1	ORECTC	carbonic dehydrat
108	15	46.9	219	2	H90677	carbonic anhydrase
109	15	46.9	219	2	D85528	carbonic anhydrase
110	15	46.9	220	2	S22493	coat protein - Cym
111	15	46.9	223	2	T33194	hypothetical prote
112	15	46.9	223	2	T33193	hypothetical prote
113	15	46.9	227	2	C90899	probable tail asse
114	15	46.9	230	2	T49555	related to high cy
115	15	46.9	231	2	S46511	isoprenyl transfer
116	15	46.9	231	2	S41894	gene 3 protein
117	15	46.9	232	1	UH0367	ribonuclease (EC 3
118	15	46.9	238	1	AB2007	transposase all160
119	15	46.9	244	1	T31838	hypothetical prote
120	15	46.9	245	1	KYBOA	chymotrypsin (EC 3
121	15	46.9	245	1	KYBOB	chymotrypsin (EC 3
122	15	46.9	247	2	B90997	probable tail asse
123	15	46.9	247	2	T27778	hypothetical prote
124	15	46.9	248	2	T03868	hypothetical prote
125	15	46.9	248	2	T03869	hypothetical prote
126	15	46.9	248	2	T31841	hypothetical prote
127	15	46.9	256	1	TRSP	trypsin-like prote
128	15	46.9	256	1	T27836	hypothetical prote
129	15	46.9	258	2	C86286	protein Pg1.16 (I
130	15	46.9	260	2	E90344	hypothetical prote
131	15	46.9	262	1	G64977	hydroxyethylthiazol
132	15	46.9	262	1	B85837	hydroxyethylthiazol
133	15	46.9	262	2	C90992	hydroxyethylthiazol
134	15	46.9	263	1	KYRTB	chymotrypsin (EC 3
135	15	46.9	263	1	KY1299	chymotrypsin (EC 3
136	15	46.9	263	2	A21195	chymotrypsin (EC 3
137	15	46.9	264	2	T38136	chymotrypsin-like
138	15	46.9	264	2	P96807	unknown protein T3
139	15	46.9	273	2	A05113	hypothetical prote
140	15	46.9	274	2	S64523	translation elonga
141	15	46.9	293	2	A85817	hypothetical prote
142	15	46.9	293	2	B85718	probable tail comp
143	15	46.9	293	2	T31840	hypothetical prote
144	15	46.9	294	2	B84640	CONSTANS-like B-do
145	15	46.9	297	2	H72779	hypothetical prote
146	15	46.9	307	2	T08125	CONSTANS protein h
147	15	46.9	307	2	T45213	superoxide dismuta
148	15	46.9	311	2	T15268	hypothetical prote
149	15	46.9	312	2	JC5522	uncoupling protei
150	15	46.9	312	2	T31834	hypothetical prote
151	15	46.9	323	2	S72915	hypothetical prote
152	15	46.9	326	2	T19996	hypothetical prote
153	15	46.9	337	2	T23794	hypothetical prote
154	15	46.9	344	2	S42383	hypothetical prote
155	15	46.9	345	2	A05279	hypothetical prote
156	15	46.9	363	2	JC5536	C-Fringe protein 1
157	15	46.9	368	2	S18758	hypothetical prote
158	15	46.9	379	2	B84073	hypothetical prote
159	15	46.9	382	2	T27058	sulfate adenyllylr
160	15	46.9	385	2	T25492	hypothetical prote
161	15	46.9	386	2	T32240	hypothetical prote
162	15	46.9	390	2	C95954	hypothetical prote
163	15	46.9	392	2	S65693	opioid receptor mu
164	15	46.9	395	1	A35629	mevalonate kinase
165	15	46.9	400	2	T56553	mu opiate receptor
166	15	46.9	410	2	T39115	formamidase-like p
167	15	46.9	412	2	AF0668	probable Benzozate
168	15	46.9	423	1	TFV2E	transforming prote
169	15	46.9	423	2	S38953	carboxypeptidase D
170	15	46.9	423	2	A41204	carboxypeptidase B
171	15	46.9	450	1	S13730	pmbA protein - Bsc
172	15	46.9	450	2	AE1056	probable PmbA prot
173	15	46.9	450	2	D86121	maturation of anti
174	15	46.9	450	2	D91280	maturation of anti
175	15	46.9	454	2	T34297	hypothetical prote
176	15	46.9	459	2	I48854	gene murine tumour
177	15	46.9	461	1	A35386	tumor necrosis fac
178	15	46.9	470	2	A40697	63k spem flagella
179	15	46.9	471	2	T39571	probable regulator
180	15	46.9	474	2	B38634	tumor necrosis fac
181	15	46.9	474	2	T43504	sp6 protein - fts
182	15	46.9	476	2	T29054	probable transmemb
183	15	46.9	482	2	S65766	G protein-coupled
184	15	46.9	488	2	T10794	phosphorin III -
185	15	46.9	493	2	S46826	hypothetical prote
186	15	46.9	493	2	S36931	chitinase (EC 3.2.
187	15	46.9	495	2	S31493	env polypeptin -
188	15	46.9	503	2	S59698	HSP1 protein - Yea
189	15	46.9	504	2	C96485	probable carnitine
190	15	46.9	504	2	C96634	probable carnitine
191	15	46.9	504	2	H64724	probable carnitine
192	15	46.9	505	2	AD0511	probable carnitine
193	15	46.9	505	2	T19971	hypothetical prote
194	15	46.9	506	2	T19973	hypothetical prote
195	15	46.9	512	2	S28267	phosphorin I prec
196	15	46.9	512	2	T06585	ammonium transport
197	15	46.9	525	2	P84933	IMP cyclohydrolase
198	15	46.9	526	2	T07082	lycopene epsilon-c
199	15	46.9	529	2	S51477	extracellular glyco
200	15	46.9	529	2	T48253	myb-like protein -
201	15	46.9	533	2	T00742	ubiquitin-binding
202	15	46.9	543	2	S65462	glucose transport
203	15	46.9	545	2	A46281	tetrahydrofolylpol
204	15	46.9	546	2	T11217	reverse transcript
205	15	46.9	549	2	TN0553	triacylglycerol 11
206	15	46.9	551	2	B88949	protein R09B5.4 (I
207	15	46.9	551	2	T01479	hypothetical prote
208	15	46.9	561	2	T27318	hypothetical prote
209	15	46.9	562	1	R8BY62	regulatory protei
210	15	46.9	568	2	T31692	hypothetical prote
211	15	46.9	571	2	T52325	pectinesterase (EC
212	15	46.9	576	2	H87414	pectinesterase (EC
213	15	46.9	584	2	S14952	TPR domain protein
214	15	46.9	588	2	T52330	pectinesterase hom
215	15	46.9	588	2	T02184	probable pectinest
216	15	46.9	595	2	T52327	pectinesterase (EC
217	15	46.9	595	2	T02183	probable pectinest
218	15	46.9	599	2	T10798	phenophorin-S - Vo
219	15	46.9	616	2	T29234	hypothetical prote
220	15	46.9	626	2	T27319	hypothetical prote
221	15	46.9	634	1	H64390	carbon-monoxide de
222	15	46.9	638	2	T27959	hypothetical prote
223	15	46.9	641	2	T05497	hypothetical prote
224	15	46.9	646	2	T35002	probable respirato
225	15	46.9	661	2	E71427	hypothetical prote
226	15	46.9	670	2	S36616	regulatory protei
227	15	46.9	676	2	T30480	envelope protein h
228	15	46.9	698	1	TFRHUP	transferrin precu
229	15	46.9	698	2	D90771	hypothetical prote
230	15	46.9	698	2	H85633	hypothetical prote
231	15	46.9	700	2	P64839	yeast protein precu
232	15	46.9	700	2	T32205	hypothetical prote
233	15	46.9	703	2	A45543	lactoferrin precu
234	15	46.9	713	2	A35502	major surface-labe
235	15	46.9	716	2	T09462	juvenile hormone r
236	15	46.9	717	2	S12100	transferrin precu
237	15	46.9	717	2	T25431	hypothetical prote
238	15	46.9	733	2	A95421	probable oxidoredu
239	15	46.9	738	2	T26421	hypothetical prote
240	15	46.9	739	2	B88553	protein K04H4.2b (
241	15	46.9	744	2	S57061	nuclear pore prote
242	15	46.9	744	2	B45046	basic juvenile hor
243	15	46.9	748	2	E95889	probable dehydroge
244	15	46.9	749	2	H91074	transcription regu
245	15	46.9	750	2	G85919	transcription regu
246	15	46.9	750	2	D65051	hyPF protein - Bsc
247	15	46.9	751	2	P87789	protein C34G6.2 (I

249	15	46.9	752	2	G85941	hypothetical prote	322	15	46.9	2090	2	T30075	hypothetical prote
250	15	46.9	752	2	C91096	hypothetical prote	323	15	46.9	2153	2	T30074	hypothetical prote
251	15	46.9	752	2	B65070	carilge oxidoredu	324	15	46.9	2201	2	A54774	ATP binding cass
252	15	46.9	755	2	A44315	gene 52 protei	325	15	46.9	2476	2	T34022	zonadhesin - pig
253	15	46.9	771	1	W2B852	furin (EC 3.4.21.7	326	15	46.9	2531	2	T31070	norch homolog - se
254	15	46.9	793	1	KXMSF	furin (EC 3.4.21.7	327	15	46.9	2844	2	S28291	hypothetical prote
255	15	46.9	793	1	KXRTF	furin (EC 3.4.21.7	328	15	46.9	2910	2	T42214	otogelin - mouse
256	15	46.9	794	1	KXHF	furin (EC 3.4.21.7	329	15	46.9	2918	2	A54105	fibillin-2 precu
257	15	46.9	797	1	T46044	furin (EC 3.4.21.7	330	15	46.9	3097	2	T00021	hypothetical prote
258	15	46.9	806	2	T18840	hypothetical prote	331	15	46.9	3147	2	T21328	hypothetical prote
259	15	46.9	814	2	T49207	receptor kinase-11	332	15	46.9	3172	2	S18253	laminin alpha-1 ch
260	15	46.9	846	1	VCLJND	env polypeptid pr	333	15	46.9	4660	2	T42737	gp330 protei
261	15	46.9	847	1	S36337	histidine decarbox	334	15	46.9	4957	2	T03455	ALR protei
262	15	46.9	849	2	T49342	fzr related protei	335	15	46.9	5262	2	T03454	sublison A [vali
263	15	46.9	852	2	S41886	DNA repair protei	336	14	43.8	43	2	A69704	fuivocin C - Myoc
264	15	46.9	853	2	S54384	envelope polypep	337	14	43.8	45	1	FRYZ	calitoxin - sea an
265	15	46.9	854	1	VCLJST	env polypeptid pr	338	14	43.8	46	2	A31863	vib7 protei
266	15	46.9	855	1	VCLJZR	env polypeptid pr	339	14	43.8	55	1	B7AGA6	lipoprotein R21 pr
267	15	46.9	856	1	A44963	env polypeptid pr	340	14	43.8	60	2	UN0750	lipoprotein R21 pr
268	15	46.9	863	2	A53034	gag polypeptid -	341	14	43.8	61	2	H90831	metallothionein-2a
269	15	46.9	863	2	T27958	hypothetical prote	342	14	43.8	62	2	S54336	hypothetical prote
270	15	46.9	873	2	B87049	conserved hypotet	343	14	43.8	62	2	S54335	hypothetical prote
271	15	46.9	888	1	GMLJHD	pol polypeptid -	344	14	43.8	62	2	T10302	hypothetical prote
272	15	46.9	890	2	T21000	hypothetical prote	345	14	43.8	65	2	S77379	hypothetical prote
273	15	46.9	891	2	E96590	hypothetical prote	346	14	43.8	73	2	F70641	hypothetical prote
274	15	46.9	895	2	B96775	hypothetical prote	347	14	43.8	73	2	G97142	hypothetical prote
275	15	46.9	911	2	S46497	aspartate kinase (	348	14	43.8	75	2	A10863	probable lipoprote
276	15	46.9	911	2	B86438	hypothetical prote	349	14	43.8	76	2	E91087	hypothetical prote
277	15	46.9	914	2	T17233	hypothetical prote	350	14	43.8	76	2	E65063	hypothetical prote
278	15	46.9	915	2	T21773	hypothetical prote	351	14	43.8	77	2	T10276	hypothetical prote
279	15	46.9	928	2	T20035	glycoprotein GP330	352	14	43.8	78	1	IKEC51	collagen V immunit
280	15	46.9	972	2	A30363	hypothetical prote	353	14	43.8	80	2	PM0044	estrogen receptor
281	15	46.9	975	2	T48107	hypothetical prote	354	14	43.8	83	2	H81188	hypothetical prote
282	15	46.9	978	2	H86319	env polypeptid -	355	14	43.8	85	2	E83008	hypothetical prote
283	15	46.9	982	1	VCLJLK	hypothetical prote	356	14	43.8	86	2	S78488	Ig kappa chain V r
284	15	46.9	995	2	S50358	DNA-directed DNA p	357	14	43.8	86	2	T07829	systemic acquired
285	15	46.9	1012	1	DJB665	DNA-directed DNA p	358	14	43.8	88	2	JC1126	major allergen cha
286	15	46.9	1012	2	T44185	probable DNA-direc	359	14	43.8	91	2	J80036	Clara cell 10K pro
287	15	46.9	1012	2	T43998	DNA polymerase [m	360	14	43.8	91	2	B87324	hypothetical prote
288	15	46.9	1012	2	T41940	DNA polymerase - h	361	14	43.8	91	2	T37327	insulin homolog ce
289	15	46.9	1013	2	T01920	DNA-directed DNA p	362	14	43.8	92	1	UCGMS	proteinasin inhibi
290	15	46.9	1014	2	T30431	DNA-directed DNA p	363	14	43.8	92	1	UCGMS	uterglobin precu
291	15	46.9	1018	2	T13693	hypothetical prote	364	14	43.8	96	2	C71614	ribosomal protei
292	15	46.9	1047	2	S19508	MSH3 protei	365	14	43.8	96	2	A36581	polychlorinated bi
293	15	46.9	1059	2	T22545	hypothetical prote	366	14	43.8	98	2	S72866	hypothetical prote
294	15	46.9	1063	2	T46284	hypothetical prote	367	14	43.8	98	2	A10334	conserved hypotet
295	15	46.9	1064	2	A40136	fibropellin 1a - s	368	14	43.8	99	2	UC2136	monocyte chemotac
296	15	46.9	1083	2	S59780	hypothetical prote	369	14	43.8	99	2	S60230	glibetellin-regula
297	15	46.9	1095	2	T13964	probable histone d	370	14	43.8	100	2	S64316	hypothetical prote
298	15	46.9	1108	2	T16875	hypothetical prote	371	14	43.8	102	2	F70534	hypothetical prote
299	15	46.9	1118	2	A48292	mucin, tracheobron	372	14	43.8	103	2	H84599	hypothetical prote
300	15	46.9	1136	2	AH1227	different proteins	373	14	43.8	103	2	S70182	hypothetical prote
301	15	46.9	1193	2	T50729	magnesium-protopo	374	14	43.8	105	2	S42212	hydroxymethylgluta
302	15	46.9	1224	2	T07446	DNA-directed RNA p	375	14	43.8	105	2	S42214	hydroxymethylgluta
303	15	46.9	1245	2	D86260	protein T12C24.22	376	14	43.8	105	2	S17345	hydroxymethylgluta
304	15	46.9	1274	2	T10729	transferrin-like p	377	14	43.8	105	2	S42213	hydroxymethylgluta
305	15	46.9	1274	2	T42017	cysteine rich prot	378	14	43.8	105	2	T42215	pectin esterase ho
306	15	46.9	1324	2	S06187	RNA2 polypeptid -	379	14	43.8	105	2	T14406	galactose-binding
307	15	46.9	1357	2	B96696	protein FIN21.4 [1	380	14	43.8	105	2	A37961	zinc finger protei
308	15	46.9	1391	2	T07406	hypothetical prote	381	14	43.8	105	2	B48827	hypothetical prote
309	15	46.9	1477	2	T18534	protein-tyrosine k	382	14	43.8	106	4	S57386	putidaredoxin [val
310	15	46.9	1524	2	T03037	polypeptid - Afri	383	14	43.8	107	1	PKPSPR	hypothetical prote
311	15	46.9	1524	2	T04464	hypothetical prote	384	14	43.8	107	2	T49527	hypothetical prote
312	15	46.9	1584	2	T00026	brain-specific ang	385	14	43.8	109	2	E72850	Acotf-5 protei
313	15	46.9	1687	2	T26318	EGF repeat cransme	386	14	43.8	109	2	T41895	orfs homolog orf13
314	15	46.9	1687	2	T30176	hypothetical prote	387	14	43.8	109	2	A13178	conserved hypotet
315	15	46.9	1722	2	B89753	protein F1IC7.4 [1	388	14	43.8	111	2	S64475	hypothetical prote
316	15	46.9	1743	2	T26859	hypothetical prote	389	14	43.8	112	2	S62929	hypothetical prote
317	15	46.9	1820	2	A55494	latent transforin	390	14	43.8	113	2	S07216	gonadotropin I bet
318	15	46.9	1827	2	A35694	cutl protein - fls	391	14	43.8	113	2	T20299	hypothetical prote
319	15	46.9	1828	2	T41455	cutl protein - fls	392	14	43.8	116	1	S15197	hyp4 protei
320	15	46.9	1863	1	A58881	breast/ovarian can	393	14	43.8	116	2	F85921	pletotropic eftec
321	15	46.9	1866	1	GNME2C	genome polypeptid	394	14	43.8	116	2	F91076	Hyp4 protei

395	14	43.8	116	2	A54598	universal minicirc
396	14	43.8	116	2	D95280	hypothetical prote
397	14	43.8	116	2	F81285	probable periplasm
398	14	43.8	118	1	PSSNM1	phospholipase A2 (
399	14	43.8	118	1	PSSNM3	phospholipase A2 (
400	14	43.8	118	2	D34860	phospholipase A2 (
401	14	43.8	118	2	E34860	phospholipase A2 (
402	14	43.8	118	2	G34860	phospholipase A2 (
403	14	43.8	118	2	F34860	phospholipase A2 (
404	14	43.8	118	2	H34860	phospholipase A2 (
405	14	43.8	118	2	AB0847	Hypa protein (impo
406	14	43.8	119	2	AC3199	transcription regu
407	14	43.8	120	2	F70424	pilin - Aquifex ae
408	14	43.8	121	2	H71351	probable ribosomal
409	14	43.8	123	2	T04635	hypothetical prote
410	14	43.8	124	2	E70754	hypothetical prote
411	14	43.8	125	2	AF2154	hypothetical prote
412	14	43.8	127	2	F70910	hypothetical prote
413	14	43.8	128	2	D69326	conserved hypothet
414	14	43.8	129	2	AH3212	hypothetical prote
415	14	43.8	131	1	BGE02	spermatid transiti
416	14	43.8	133	2	AB1108	gonadotropin I bet
417	14	43.8	137	2	B36179	gonadotropin I bet
418	14	43.8	137	2	IS1231	gonadotropin I bet
419	14	43.8	137	2	S34349	gonadotropin I bet
420	14	43.8	137	2	S55364	serine proteinase
421	14	43.8	138	2	T21792	hypothetical prote
422	14	43.8	139	2	S69458	hypothetical prote
423	14	43.8	139	2	PC4217	hypothetical 139 p
424	14	43.8	139	2	T12618	homobox protein H
425	14	43.8	140	2	T04904	hypothetical prote
426	14	43.8	141	2	A85994	probable transcrip
427	14	43.8	141	2	AB1009	probable Zn(II)-re
428	14	43.8	141	2	E91148	zinc (II) responsi
429	14	43.8	141	2	I67892	embryonic abundan
430	14	43.8	142	2	T09251	probable CO-induce
431	14	43.8	142	2	T51316	ORF MSV132 probabl
432	14	43.8	142	2	T28293	hypothetical prote
433	14	43.8	142	2	B84040	hypothetical prote
434	14	43.8	142	2	US0510	fusaric acid resis
435	14	43.8	142	2	T03976	hypothetical prote
436	14	43.8	142	4	S13768	MHC class I histoc
437	14	43.8	143	2	H83935	hypothetical prote
438	14	43.8	143	2	T00293	hypothetical prote
439	14	43.8	143	2	T27692	hypothetical prote
440	14	43.8	144	2	S01391	phospholipase A2 (
441	14	43.8	144	2	T05641	hypothetical prote
442	14	43.8	146	2	T25356	hypothetical prote
443	14	43.8	146	2	C86187	YUPH12.12 (import
444	14	43.8	146	2	AD3427	transposase BME114
445	14	43.8	147	2	S42552	proline-rich prote
446	14	43.8	148	2	T30627	hypothetical prote
447	14	43.8	148	2	T48981	hypothetical prote
448	14	43.8	148	2	B72513	hypothetical prote
449	14	43.8	150	2	S11441	plasma C precursor
450	14	43.8	151	2	E83550	positive regulator
451	14	43.8	151	2	T09256	heat shock protein
452	14	43.8	152	1	KRSHHC	keratin high-sulfu
453	14	43.8	152	2	I47111	high-sulfur wool m
454	14	43.8	152	2	I47109	high-sulfur wool m
455	14	43.8	152	2	I47108	high-sulfur wool m
456	14	43.8	152	2	I47112	high-sulfur wool m
457	14	43.8	152	2	A84303	hypothetical prote
458	14	43.8	152	2	T09253	heat shock protein
459	14	43.8	152	2	D89753	protein F11C7.2 (l
460	14	43.8	153	2	F75514	conserved hypothet
461	14	43.8	153	2	T32950	hypothetical prote
462	14	43.8	154	1	GOVLAV	gene X protein - h
463	14	43.8	154	1	QOVLCP	gene X protein - h
464	14	43.8	154	1	QOVLKS	gene X protein - h
465	14	43.8	154	2	S47408	gene X protein - h
466	14	43.8	154	2	S20756	gene X protein - h
467	14	43.8	154	2	S67503	gene X protein - h
468	14	43.8	154	2	S12542	gene X protein - h
469	14	43.8	154	2	S33687	gene X protein - h
470	14	43.8	155	2	D72761	hypothetical prote
471	14	43.8	156	2	I67751	dopamine receptor
472	14	43.8	156	2	B83185	probable transcrip
473	14	43.8	156	2	B82073	sigma-E factor reg
474	14	43.8	159	1	B83299	sigma-E factor reg
475	14	43.8	159	2	AP0829	sigma-E factor reg
476	14	43.8	159	2	B85903	sigma-E factor, ne
477	14	43.8	159	2	D91058	sigma-E factor reg
478	14	43.8	160	2	T02999	sigma-E factor reg
479	14	43.8	160	2	T28953	hypothetical prote
480	14	43.8	160	2	JC7606	globin chain al pr
481	14	43.8	161	2	T21405	hypothetical prote
482	14	43.8	162	2	A87755	protein T21B12.3 (
483	14	43.8	166	2	E71342	probable deoxycytl
484	14	43.8	167	2	D81049	hypothetical prote
485	14	43.8	171	2	S09903	hypothetical prote
486	14	43.8	171	2	S09759	hypothetical prote
487	14	43.8	171	2	H83360	probable acetyltra
488	14	43.8	172	2	B81828	hypothetical integ
489	14	43.8	173	2	AB3450	invasion membrane
490	14	43.8	174	2	AB1923	probable membrane
491	14	43.8	174	1	JQ1625	small hydrophobic
492	14	43.8	176	2	T28762	hypothetical prote
493	14	43.8	177	2	B81085	conserved hypothet
494	14	43.8	177	2	S33166	protective antigen
495	14	43.8	177	2	H95981	conserved hypothet
496	14	43.8	178	2	A81687	conserved hypothet
497	14	43.8	178	2	S64257	hypothetical prote
498	14	43.8	181	2	S78572	EST3 protein - yea
499	14	43.8	182	2	T17625	hypothetical prote
500	14	43.8	184	2	D83234	hypothetical prote
501	14	43.8	185	2	T29777	hypothetical prote
502	14	43.8	187	2	T09390	21K protein precur
503	14	43.8	189	2	T28554	hypothetical prote
504	14	43.8	189	2	D36849	Al2L protein - var
505	14	43.8	189	2	B72165	Al3L protein - var
506	14	43.8	189	2	G96514	hypothetical prote
507	14	43.8	189	2	UX0235	core protein MGC-2
508	14	43.8	191	2	JQ1999	hypothetical 21.5K
509	14	43.8	191	2	T32278	hypothetical prote
510	14	43.8	192	2	E42518	Al2L protein - vac
511	14	43.8	193	2	D83296	hypothetical prote
512	14	43.8	196	2	T26943	hypothetical prote
513	14	43.8	197	2	B85068	hypothetical prote
514	14	43.8	197	2	T26829	hypothetical prote
515	14	43.8	197	2	C88809	protein Y43CSB.1 (
516	14	43.8	197	2	C86391	hypothetical prote
517	14	43.8	198	2	S72709	Leb1170 C3 229 pr
518	14	43.8	198	2	A36646	ORF1 protein - equ
519	14	43.8	200	2	B83642	hypothetical prote
520	14	43.8	201	2	S38652	resolvasse - Klebsi
521	14	43.8	201	2	D82779	hypothetical prote
522	14	43.8	203	2	E70150	hypothetical prote
523	14	43.8	204	2	T02386	hypothetical prote
524	14	43.8	205	2	T02385	hypothetical prote
525	14	43.8	205	2	T04918	hypothetical prote
526	14	43.8	207	2	T22436	hypothetical prote
527	14	43.8	207	2	S70533	bbk2.10 protein pr
528	14	43.8	208	2	B75458	hypothetical prote
529	14	43.8	208	2	P00278	hypothetical prote
530	14	43.8	209	2	C70857	probable ligpa prot
531	14	43.8	211	2	S25938	env protein - huma
532	14	43.8	211	2	T18799	hypothetical prote
533	14	43.8	212	2	S09623	agglutinin isolect
534	14	43.8	213	1	A6EWT2	agglutinin isolect
535	14	43.8	213	1	T27841	hypothetical prote
536	14	43.8	213	2	B82698	conserved hypothet
537	14	43.8	215	2	A60166	hemiferrin - bovin
538	14	43.8	215	2	A71541	hypothetical prote
539	14	43.8	216	1	A39684	hemiferrin - rat
540	14	43.8	216	2	H88102	protein W10G11.13

541	14	43.8	216	2	S60048	chlorophyll a/c-bi	614	14	43.8	273	2	G83403	conserved hypothet
542	14	43.8	217	2	T25618	hypothetical prote	615	14	43.8	274	2	JC4171	trypcase (EC 3.4.2
543	14	43.8	218	2	D69428	hypothetical prote	616	14	43.8	275	2	AE3112	transcription regu
544	14	43.8	219	2	T34373	hypothetical prote	617	14	43.8	276	2	H98174	transcription regu
545	14	43.8	219	2	S25939	env. proteain - huma	618	14	43.8	277	2	A33410	trypcase (EC 3.4.2
546	14	43.8	220	2	D83390	carbonate dehydrat	619	14	43.8	275	2	G83522	conserved hypothet
547	14	43.8	222	1	MWVZB4	antithrombin-III h	620	14	43.8	275	2	T29952	hypothetical prote
548	14	43.8	223	2	A65172	hypothetical prote	621	14	43.8	277	2	S71222	xyloglucan endo-1,
549	14	43.8	223	2	F89567	protein T08A9.3 [l	622	14	43.8	280	2	D97185	dtDP-4-dehydroham
550	14	43.8	223	2	T37974	probable peroxisom	623	14	43.8	280	2	G01884	LIM protein FHL-1,
551	14	43.8	224	2	T01325	hypothetical prote	624	14	43.8	280	2	F95998	probable sugar nuc
552	14	43.8	225	2	C86563	CT465 hypothetical	625	14	43.8	281	2	AE2248	phosphomethylpyrim
553	14	43.8	225	2	C72062	conserved hypothet	626	14	43.8	281	2	C86221	hypothetical prote
554	14	43.8	226	2	T27843	hypothetical prote	627	14	43.8	282	2	T15556	hypothetical prote
555	14	43.8	226	2	T20219	hypothetical prote	628	14	43.8	282	2	T15640	hypothetical prote
556	14	43.8	229	2	T27840	hypothetical prote	629	14	43.8	283	2	T18969	5'-methylthiodeno
557	14	43.8	229	2	B82544	hypothetical prote	630	14	43.8	283	2	T21622	hypothetical prote
558	14	43.8	231	2	S28703	hypothetical prote	631	14	43.8	284	2	H97220	transcriptional reg
559	14	43.8	231	2	F64642	hypothetical prote	632	14	43.8	284	2	T06200	xyloglucan endo-1,
560	14	43.8	231	2	T24669	hypothetical prote	633	14	43.8	284	2	S04278	hypoxanthine phosp
561	14	43.8	232	2	E71325	probable V-type AT	634	14	43.8	285	1	IDECRP	replication initia
562	14	43.8	234	2	T32843	hypothetical prote	635	14	43.8	285	1	I64780	replication-associ
563	14	43.8	237	2	T28070	hypothetical prote	636	14	43.8	285	2	T00306	replication initia
564	14	43.8	238	2	T04166	chamaatin-like pro	637	14	43.8	285	2	B88546	protein R107.2 [lm
565	14	43.8	238	2	T40568	hypothetical prote	638	14	43.8	285	2	T24665	hypothetical prote
566	14	43.8	239	2	B81345	probable periplasm	639	14	43.8	285	2	S09614	hypoxanthine phosp
567	14	43.8	240	2	T20319	hypothetical prote	640	14	43.8	286	2	S30872	ethanolamine ammon
568	14	43.8	242	2	T29854	hypothetical prote	641	14	43.8	286	2	B36570	hypothetical prote
569	14	43.8	243	2	A33329	hypothetical prote	642	14	43.8	286	2	T23354	hypothetical prote
570	14	43.8	243	2	J80204	testis-specific pr	643	14	43.8	286	2	T31494	hypothetical prote
571	14	43.8	244	1	JC2018	phosducin - mouse	644	14	43.8	287	2	A85059	hypothetical prote
572	14	43.8	244	1	T37493	puarative oxidore	645	14	43.8	288	2	T02546	hypothetical prote
573	14	43.8	245	2	S52096	phosducin - cat	646	14	43.8	288	2	T21790	hypothetical prote
574	14	43.8	245	2	AF0412	sugar fermentation	647	14	43.8	289	2	T34688	probable lipoprote
575	14	43.8	245	2	T24565	hypothetical prote	648	14	43.8	290	2	C71321	hypothetical prote
576	14	43.8	246	1	A35422	phosducin, retinal	649	14	43.8	290	2	T10104	maturase-like prote
577	14	43.8	247	2	T43324	synaptogyrin homol	650	14	43.8	292	2	T19669	hypothetical prote
578	14	43.8	248	2	S49323	chymotrypsin (EC 3	651	14	43.8	292	2	G88783	protein C33A12.16
579	14	43.8	248	2	T21786	hypothetical prote	652	14	43.8	293	2	E49539	xyloglucan endo-1,
580	14	43.8	248	2	T23769	hypothetical prote	653	14	43.8	294	2	A48844	TGF alpha-like pro
581	14	43.8	249	2	T31837	hypothetical prote	654	14	43.8	295	2	G65018	ethanolamine ammon
582	14	43.8	249	2	T21785	hypothetical prote	655	14	43.8	295	2	B85887	ethanolamine ammon
583	14	43.8	250	2	T31836	hypothetical prote	656	14	43.8	295	2	G91042	ethanolamine ammon
584	14	43.8	250	2	T31835	hypothetical prote	657	14	43.8	295	2	C90740	probable transmemb
585	14	43.8	251	2	T14407	pectin esterase ho	658	14	43.8	295	2	A46017	zinc finger protei
586	14	43.8	251	2	A55523	hypothetical prote	659	14	43.8	295	2	E95590	probable transmemb
587	14	43.8	251	2	B39685	GRESAG protein 2.1	660	14	43.8	295	2	B64818	probable membrane
588	14	43.8	252	2	AB0213	conserved hypothet	661	14	43.8	296	2	H87511	hypothetical prote
589	14	43.8	252	2	S61178	hypothetical prote	662	14	43.8	297	2	T18660	hypothetical prote
590	14	43.8	253	2	T04059	hypothetical prote	663	14	43.8	297	2	A98294	hypothetical prote
591	14	43.8	253	2	E89130	protein F52E1.2 [l	664	14	43.8	297	2	AH2989	permease (imported
592	14	43.8	254	2	AB1304	dihydroorotate deh	665	14	43.8	297	2	S06267	surface antigen H
593	14	43.8	254	2	AB1676	sugar fermentation	666	14	43.8	298	2	AF0813	ethanolamine ammon
594	14	43.8	256	2	F82304	hypothetical prote	667	14	43.8	299	2	T43989	minor capsid prote
595	14	43.8	257	2	T28946	hypothetical prote	668	14	43.8	300	2	C88970	transcription regu
596	14	43.8	257	2	T47491	hypothetical prote	669	14	43.8	301	2	F82287	transcription regu
597	14	43.8	258	2	B82249	uridine phosphoryl	670	14	43.8	301	2	B84325	tRNA-pseudouridine
598	14	43.8	259	1	IOHUI	insulin-like growt	671	14	43.8	302	2	D65054	hypothetical prote
599	14	43.8	259	2	T27783	hypothetical prote	672	14	43.8	302	2	H84731	hypothetical prote
600	14	43.8	259	2	E84427	hypothetical prote	673	14	43.8	302	2	JC2071	chitinase (EC 3.2.
601	14	43.8	262	2	T05084	hypothetical prote	674	14	43.8	304	2	A72596	hypothetical prote
602	14	43.8	263	2	S44668	hypothetical prote	675	14	43.8	305	2	T28005	hypothetical prote
603	14	43.8	264	2	T09324	ZK370.7 protein -	676	14	43.8	305	2	S32834	methylviologen-red
604	14	43.8	264	2	AE2274	capsid protein - h	677	14	43.8	310	2	T41985	hypothetical prote
605	14	43.8	265	2	S62364	D-2-chloropropion	678	14	43.8	311	2	A28446	transferrin - mous
606	14	43.8	267	2	C81651	conserved hypothet	679	14	43.8	311	2	JC7873	L-thiamose-binding
607	14	43.8	267	2	H86548	polymorphic outer	680	14	43.8	312	2	T09443	zinc metalloprotei
608	14	43.8	270	2	I37278	complement factor	681	14	43.8	312	2	E82276	hypothetical prote
609	14	43.8	270	2	C96731	unknown protein F5	682	14	43.8	312	2	T17969	hypothetical prote
610	14	43.8	270	2	F89632	protein F13B6.3 [i	683	14	43.8	312	2	T45061	hypothetical prote
611	14	43.8	271	2	T40238	hypothetical prote	684	14	43.8	314	2	T28879	hypothetical prote
612	14	43.8	272	2	T15351	hypothetical prote	685	14	43.8	317	2	A33985	wound-inducible ch
613	14	43.8	272	2	T21789	hypothetical prote	686	14	43.8	318	2	T04403	probable chitinase

687	14	43.8	319	2	H83253	conserved hypochet	760	14	43.8	369	2	S60186	peridinin-chloroph
688	14	43.8	319	2	B86395	probable origin re	761	14	43.8	370	2	S15013	wnt-1 protein - ze
689	14	43.8	320	2	S38670	chitinase (EC 3.2.	762	14	43.8	372	1	W2ML31	E2 protein - human
690	14	43.8	320	2	T28992	hypothetical prote	763	14	43.8	372	2	T31060	hypothetical prote
691	14	43.8	322	1	A29928	membrane-associate	764	14	43.8	372	2	T45410	hypothetical prote
692	14	43.8	322	2	H87131	dieneolactone hydrol	765	14	43.8	372	2	T29359	hypothetical prote
693	14	43.8	323	2	G87358	hypothetical prote	766	14	43.8	373	2	A70856	probable lipz prot
694	14	43.8	323	2	T32279	hypothetical prote	767	14	43.8	375	2	T17488	hypothetical prote
695	14	43.8	324	2	A47172	transforming growt	768	14	43.8	375	2	G02300	pancreatic polypep
696	14	43.8	324	2	JC2395	fas antigen precur	769	14	43.8	375	2	I39182	neuropeptide y rec
697	14	43.8	325	2	I65354	stem cell antigen	770	14	43.8	375	2	S63685	neuropeptide y rec
698	14	43.8	327	2	P96814	hypothetical prote	771	14	43.8	375	2	S52143	amc protein - brw
699	14	43.8	328	2	G89152	protein C24B5.5 f	772	14	43.8	376	2	A96720	hypothetical prote
700	14	43.8	328	2	T22636	hypothetical prote	773	14	43.8	378	2	B64044	rod shape-determ
701	14	43.8	329	2	T18619	hypothetical prote	774	14	43.8	379	2	G82380	hypothetical prote
702	14	43.8	331	2	T06227	peroxidase (EC 1.1	775	14	43.8	380	2	C96754	Similar to part of
703	14	43.8	331	2	C71468	probable glycerol-	776	14	43.8	381	2	G83835	sulfate adenylyl tr
704	14	43.8	331	2	B83486	cobalamin biosynth	777	14	43.8	382	2	T27076	hypothetical prote
705	14	43.8	332	2	T04484	probable chitinase	778	14	43.8	382	2	B46131	neuropeptide y pep
706	14	43.8	332	2	T21458	hypothetical prote	779	14	43.8	382	2	S27388	neuropeptide y rec
707	14	43.8	332	2	A26663	nautilus protein -	780	14	43.8	382	2	I52565	stem cell antigen
708	14	43.8	333	2	D88970	protein f15B1.11	781	14	43.8	382	2	B91083	hypothetical membr
709	14	43.8	333	2	T34498	hypothetical prote	782	14	43.8	382	2	F85928	hypothetical prote
710	14	43.8	334	2	B81272	probable sugar mnc	783	14	43.8	383	1	VGBEKG	glycoprotein precu
711	14	43.8	336	2	T15727	hypothetical prote	784	14	43.8	384	2	A45490	neuropeptide y pep
712	14	43.8	340	2	JC7695	G protein-coupled	785	14	43.8	384	2	S64735	retrovirus-related
713	14	43.8	341	1	WMVZHI	antithrombin-III h	786	14	43.8	384	2	T38544	probable exopolypn
714	14	43.8	341	1	H89980	conserved hypochet	787	14	43.8	385	2	S68780	probable D1-like y
715	14	43.8	344	2	E72173	D2R protein - vari	788	14	43.8	385	2	S53718	homeotic protein d
716	14	43.8	344	2	S46871	B13 protein - vari	789	14	43.8	385	2	A54785	preadipocyte facto
717	14	43.8	344	2	T28611	hypothetical prote	790	14	43.8	386	2	S72168	dopamine receptor
718	14	43.8	344	2	A70771	hypothetical prote	791	14	43.8	386	2	T12527	hypothetical prote
719	14	43.8	345	1	WMVZW2	antithrombin-III h	792	14	43.8	386	2	T19935	hypothetical prote
720	14	43.8	345	1	JN0465	apolipoprotein H p	793	14	43.8	387	2	B71611	hypothetical prote
721	14	43.8	345	2	T32018	hypothetical prote	794	14	43.8	388	2	S51316	prostaglandin E re
722	14	43.8	346	2	AP1820	siatoglycoproteina	795	14	43.8	388	2	I38750	hypothetical 43.6K
723	14	43.8	346	2	T45748	hypothetical prote	796	14	43.8	390	2	S56560	hypothetical prote
724	14	43.8	346	2	T47411	hypothetical prote	797	14	43.8	390	2	B91291	hypothetical prote
725	14	43.8	347	2	S60428	hypothetical prote	798	14	43.8	390	2	D86132	hypothetical prote
726	14	43.8	347	2	JQ2359	wheat aluminum ind	799	14	43.8	392	2	A60777	keratin 2, type I,
727	14	43.8	348	2	S64041	probable membrane	800	14	43.8	392	2	T33505	hypothetical prote
728	14	43.8	348	2	S75548	siatoglycoproteina	801	14	43.8	395	2	S58157	hypothetical prote
729	14	43.8	348	2	S19873	single-stranded DN	802	14	43.8	396	1	A42919	mevalonate kinase
730	14	43.8	349	2	AP3401	enoyl-CoA hydratase	803	14	43.8	397	2	T08345	hypothetical prote
731	14	43.8	349	2	S12863	G protein-coupled	804	14	43.8	398	2	AH0482	conserved hypochet
732	14	43.8	352	2	C96643	hypothetical prote	805	14	43.8	398	2	T19421	hypothetical prote
733	14	43.8	352	2	S77448	hypothetical prote	806	14	43.8	400	2	S32804	beta-3-adrenergic
734	14	43.8	354	2	T26599	hypothetical prote	807	14	43.8	400	2	A41679	beta-3-adrenergic
735	14	43.8	354	2	G86276	hypothetical prote	808	14	43.8	400	2	A53281	beta-3-adrenergic
736	14	43.8	356	2	A25918	thrombomodulin - b	809	14	43.8	400	2	D89810	conserved hypochet
737	14	43.8	356	2	T18590	hypothetical prote	810	14	43.8	401	2	B96570	hypothetical prote
738	14	43.8	357	2	A71295	probable alanine r	811	14	43.8	402	2	T39591	cyclin - fibron y
739	14	43.8	357	2	A97697	hypothetical prote	812	14	43.8	403	2	C82228	probable trypsin V
740	14	43.8	358	2	D84901	hypothetical prote	813	14	43.8	403	2	A31847	hypothetical prote
741	14	43.8	358	2	T23802	hypothetical prote	814	14	43.8	403	2	T49480	hypothetical prote
742	14	43.8	359	2	AD3520	galactoside transp	815	14	43.8	405	2	S65459	beta-3-adrenergic
743	14	43.8	361	2	T25305	hypothetical prote	816	14	43.8	405	2	B35340	alpha(1,3)-fucoseyl
744	14	43.8	361	1	C65132	hypothetical 38.6	817	14	43.8	406	2	A43850	glutamate N-acetyl
745	14	43.8	362	1	LPEBCRA	rare lipoprotein A	818	14	43.8	406	2	C81017	amino-acid N-acetyl
746	14	43.8	362	2	G85563	a minor lipoprotei	819	14	43.8	407	2	T19895	hypothetical prote
747	14	43.8	362	2	G90712	hypothetical prote	820	14	43.8	408	1	QRHUBR	beta-3-adrenergic
748	14	43.8	362	2	T20222	dopamine D1 recep	821	14	43.8	408	1	QRHUBR	beta-3-adrenergic
749	14	43.8	363	2	I50475	triacylglycerol 11	822	14	43.8	408	2	B93333	immediate-early pr
750	14	43.8	363	2	A34959	probable acyl-CoA	823	14	43.8	408	2	A72869	probable MPS trans
751	14	43.8	363	2	AB1134	prostaglandin E re	824	14	43.8	410	2	B84144	early gene transac
752	14	43.8	363	2	C81888	collagen alpha 2(I)	825	14	43.8	412	2	B81176	hypothetical prote
753	14	43.8	364	2	JC2115	neuropeptide Y pep	826	14	43.8	412	2	T26056	hypothetical prote
754	14	43.8	365	2	S65009	GTP-binding protei	827	14	43.8	413	1	QRHUB2	beta-2-adrenergic
755	14	43.8	366	1	S71152	peccate lyase (EC	828	14	43.8	413	2	S47527	extracellular sucr
756	14	43.8	367	1	I51426	cinnamoyl CoA redu	829	14	43.8	414	1	QRHUB3	beta-fructofuranos
757	14	43.8	367	1	S15716		830	14	43.8	414	1	QRHUB3	beta-3-adrenergic
758	14	43.8	367	1	S15716		831	14	43.8	415	2	I53040	beta-2 adrenergic
759	14	43.8	368	2	T46643		832	14	43.8	415	2	A64299	coenzyme F420 hydr





979 14 43.8 510 1 S15620 l1 protein - human  
 980 14 43.8 510 1 S15627 l1 protein - human  
 981 14 43.8 511 2 S44275 dopamine receptor  
 982 14 43.8 511 2 JC7682 spermatogenesis as  
 983 14 43.8 513 2 T38866 conserved hypothet  
 984 14 43.8 515 2 D71319 probable glucose-6  
 985 14 43.8 516 2 T33269 hypothetical prote  
 986 14 43.8 516 2 S40996 hypothetical prote  
 987 14 43.8 517 2 A71707 lipopolysaccharide  
 988 14 43.8 519 2 S78089 G-protein signalin  
 989 14 43.8 520 2 T23545 hypothetical prote  
 990 14 43.8 522 2 S71821 probable interleuk  
 991 14 43.8 523 1 S61713 carboxypeptidase C  
 992 14 43.8 524 1 P1WL58 l1 protein - human  
 993 14 43.8 527 1 SAHUP saposin precursor  
 994 14 43.8 528 2 B88989 protein F02C9.2 [l  
 995 14 43.8 528 2 T33527 hypothetical prote  
 996 14 43.8 529 2 S46116 probable regulator  
 997 14 43.8 530 2 JCS939 estrogen receptor  
 998 14 43.8 531 2 T32276 hypothetical prote  
 999 14 43.8 532 2 S18448 variant surface g1  
 1000 14 43.8 533 2 T31002 hypothetical prote

## ALIGNMENTS

## RESULT 1

SS9448 hypothetical protein YMR206w - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein YMR325.07

C/Species: Saccharomyces cerevisiae

C/Date: 30-Nov-1995 #sequence\_revision 23-Feb-1996 #text\_change 09-Jul-2004

C/Accession: S59448

R/Odell, C.; Bowman, S.

submitted to the EMBL Data Library, March 1995

A/Reference number: S59441

A/Accession: S59448

A/Molecule type: DNA

A/Residues: 1-313 <ODE>

A/Cross-references: UNIPROT:Q03695; UNIPARC:UPI000013B90A; EMBL:Z48755; NID:G736296; PID

C/Genetics:

A/Gene: MIPS:YMR206w

A/Cross-references: SGD:S0004819

A/Map position: 13R

Query Match 56.2%; Score 18; DB 2; Length 313;  
 Best Local Similarity 20.0%; Pred. No. 0.44; Mismatches 8; Indels 0; Gaps 0;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

## RESULT 2

E75433 hypothetical protein - Deinococcus radiodurans (strain R1)

C/Species: Deinococcus radiodurans

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C/Accession: E75433

R/White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A/Reference number: A75250; MUID:20036896; PMID:10567266

A/Accession: E75433

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-480 <WHI>  
 A/Cross-references: UNIPROT:Q9RV87; UNIPARC:UPI00000C18A8; GB:AE001963; GB:AE005513; NID

A/Experimental source: strain R1  
 C/Genetics:  
 A/Gene: DR1142  
 A/Map position: 1  
 C/Superfamily: Deinococcus radiodurans hypothetical protein DR1142

Query Match 56.2%; Score 18; DB 2; Length 480;  
 Best Local Similarity 20.0%; Pred. No. 0.51; Mismatches 8; Indels 0; Gaps 0;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
 DB 464 CSAATATAC 473

## RESULT 3

JQ1560 hypothetical 20.6K protein - Lymantia dispar nuclear polyhedrosis virus

N/Alternate names: hypothetical protein 4

C/Species: Lymantia dispar nuclear polyhedrosis virus, LdNPV

C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004

C/Accession: JQ1560

R./Bjornson, R.M.; Rohmann, G.F.

J. Gen. Virol. 73, 1499-1504, 1992

A/Title: Nucleotide sequence of the polyhedron envelope protein gene region of the Lyman

A/Reference number: PQ0339; MUID:92300345; PMID:1607868

A/Accession: JQ1560

A/Molecule type: DNA

A/Residues: 1-194 <BJO>

A/Cross-references: UNIPROT:P36868; UNIPARC:UPI000017A7BF; DBJ:DI0836

Query Match 53.1%; Score 17; DB 2; Length 194;  
 Best Local Similarity 20.0%; Pred. No. 1.6; Mismatches 8; Indels 0; Gaps 0;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
 DB 74 CSAATSSSC 83

## RESULT 4

JC4082 coat protein - Cymbidium mosaic virus

C/Species: Cymbidium mosaic virus

C/Date: 21-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004

C/Accession: JC4082

R./Ryu, K.H.; Yoon, K.E.; Park, W.M.

Gene 156, 303-304, 1995

A/Title: Cloning and sequencing of a cDNA encoding the coat protein of a Korean isolate

A/Reference number: JC4082; MUID:95278762; PMID:7758973

A/Accession: JC4082

A/Molecule type: mRNA

A/Residues: 1-220 <RYU>

A/Cross-references: UNIPROT:Q66152; UNIPARC:UPI0000F22DE; EMBL:X81051; NID:G897718; PID

C/Note: The authors translated the codon GTG for residue 161 as Leu

C/Superfamily: potato virus coat protein

C/keywords: coat protein

QY 4 CXXXXXXXXXC 13  
 DB 166 CSAATLTATC 175

## RESULT 5

AB2920 cobalamin biosynthetic protein [imported] - Agrobacterium tumefaciens (strain C58, Dupor

C/Species: Agrobacterium tumefaciens

C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004

C/Accession: AB2920

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I. erege, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McCelli  
A:Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, B.W.  
A>Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A:Reference number: AB2577, MUID:21608550; PMID:11743193  
A:Accession: AB2920  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-376 <KUR>  
A:Cross-references: UNIPROT:Q8UBQ6; UNIPARC:UPI0000164772; GB:AE008688; PIDD:AAI43776.1;  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: cbid  
A:Map position: circular chromosome  
C:Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 53.1%; Score 17; DB 2; Length 376;  
Best Local Similarity 20.0%; Pred. No. 2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXX 13  
DB 17 CAAATKAC 26

RESULT 6  
C97694  
cobalamin biosynthetic protein cbid (PA2908) [imported] - *Agrobacterium tumefaciens* (str  
C:Species: *Agrobacterium tumefaciens*  
C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: C97694  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Woliam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: C97694  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-388 <KUR>  
A:Cross-references: UNIPROT:Q8UBQ6; UNIPARC:UPI000002014; GB:AE007869; PIDD:AAK8508.1;  
C:Genetics:  
A:Gene: AGR\_C\_5073  
A:Map position: circular chromosome  
C:Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 53.1%; Score 17; DB 2; Length 388;  
Best Local Similarity 20.0%; Pred. No. 2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXX 13  
DB 29 CAAATKAC 38

RESULT 7  
T23167  
hypothetical protein K01C8.2 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T23167  
R:Sims, M.  
submitted to the EMBL Data Library, April 1995  
A:Reference number: Z19702  
A:Accession: T23167  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-389 <MTL>  
A:Cross-references: UNIPROT:Q21081; UNIPARC:UPI0000080155; EMBL:Z49066; PIDD:CAA8885.1;  
A:Experimental source: clone K01C8

C:Genetics:  
A:Gene: CESP:K01C8.2  
A:Map position: 2  
A:Introns: 54/2; 146/3; 208/3; 283/1; 379/3

Query Match 53.1%; Score 17; DB 2; Length 389;  
Best Local Similarity 20.0%; Pred. No. 2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXX 13  
DB 123 CSSSISTSSC 132

RESULT 8  
T37314  
probable kexin (EC 3.4.21.61) - *Caenorhabditis elegans* (fragment)  
N:Alternate names: Bisterase 4  
C:Species: *Caenorhabditis elegans*  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jul-2000  
C:Accession: T37314  
R:Thacker, C.; Peters, K.; Strayko, M.; Rose, A.M.  
Genes Dev. 9, 956-971, 1995  
A>Title: The bli-4 locus of *Caenorhabditis elegans* encodes structurally distinct kex2/b  
A:Reference number: Z21679; MUID:95293228; PMID:7774813  
A:Accession: T37314  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-570 <THA>  
A:Cross-references: UNIPARC:UPI000016B8E1; EMBL:129440; NID:G459702; PIDD:AAA98752.1; P  
C:Genetics:  
A:Gene: bli-4  
A:Map position: I  
C:Keyword: alternative splicing; hydrolase; serine proteinase

Query Match 53.1%; Score 17; DB 2; Length 570;  
Best Local Similarity 20.0%; Pred. No. 2.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXX 13  
DB 311 CTSSSANTSC 320

RESULT 9  
S33575  
dnaK-type molecular chaperone precursor, mitochondrial - *Leishmania major*  
N:Alternate names: heat shock protein 70-related protein; mitochondrial stress protein  
C:Species: *Leishmania major*  
C>Date: 09-Dec-1997 #sequence\_revision 09-Dec-1997 #text\_change 31-Dec-2004  
C:Accession: S33575; S78090; S05438  
R:Searle, S.; McCrossan, M.V.; Smith, D.F.  
J. Cell Sci. 104, 1091-1100, 1993  
A>Title: Expression of a mitochondrial stress protein in the protozoan parasite *Leishman*  
A:Reference number: S33575; MUID:93300981; PMID:8314893  
A:Accession: S33575  
A:Molecule type: DNA  
A:Residues: 1-634 <SEA>  
A:Cross-references: UNIPROT:P12076; UNIPARC:UPI0000177D24; EMBL:X64137  
R:Smith, D.F.  
submitted to the EMBL Data Library, January 1992  
A:Reference number: S78090  
A:Accession: S78090  
A:Molecule type: DNA  
A:Residues: 1-460,481-500, 'QSERETASRNQIRGEHLSG', 501-634 <SMI>  
A:Cross-references: UNIPARC:UPI000012C5C5; EMBL:X64137; NID:g311289; PIDD:CAA45498.1; E  
Nucleic Acids Res. 17, 5081-5095, 1989  
A>Title: A family of heat shock protein 70-related genes are expressed in the promastig  
A:Reference number: S05438; MUID:89345072; PMID:2762121  
A:Accession: S05438  
A>Status: translation not shown  
A:Molecule type: DNA

A:Residues: 1-249 <SEF>  
A:Crosss-references: UNIPARC:UPI000016BPLC; EMBL:X14574; NID:g9561; PIDN:CAA32713.1; PID:  
C:Genetics:  
A:Gene: lmbp70.1  
A:Genome: nuclear  
C:Function:  
A:Description: involved in protein folding and assembling/disassembling of protein comp  
C:Superfamily: bcr protein  
C:Keywords: ATP; mitochondrion; molecular chaperone; stress-induced protein  
F:1-23/Domain: transit peptide (mitochondrion) #status predicted <TNP>  
F:24-634/Product: heat shock protein 70, mitochondrial #status predicted <MAT>

Query Match 53.1%; Score 17; DB 2; Length 634;  
Best Local Similarity 20.0%; Pred. No. 2.4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
DB 7 CGSAAASAAAC 16

RESULT 10  
D87803  
protein bli-4D [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 21-Jun-2002  
C:Accession: D87803  
R:Anonymous, The C. elegans Sequencing Consortium.  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: D87803  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-942 <STO>  
A:Crosss-references: UNIPARC:UPI000016B639; GB:chr\_I; PIDN:AAB96754.1; PID:g2773243; GSPI  
C:Genetics:  
A:Gene: bli-4D  
A:Map position: 1  
C:Superfamily: kexin; subtilisin homology

Query Match 53.1%; Score 17; DB 2; Length 942;  
Best Local Similarity 20.0%; Pred. No. 2.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
DB 683 CTSSSATATSC 692

RESULT 11  
S09118  
G surface protein 168 - Paramoecium primaurelia  
C:Species: Paramoecium primaurelia  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: S09118  
R:Prat, A.  
J: Mol. Biol. 211, 521-535, 1990  
A:Title: Conserved sequences flank variable tandem repeats in two alleles of the G surf  
A:Reference number: S09118; MUID:90172419; PMID:2308165  
A:Accession: S09118  
A:Molecule type: DNA  
A:Residues: 1-2704 <PRA>  
A:Cross-references: UNIPROT:P17053; UNIPARC:UPI000012AE23; EMBL:X52133; NID:g10049; PIDN  
C:Genetics:  
A:Genetic code: SGCS  
C:Superfamily: G surface protein

Query Match 53.1%; Score 17; DB 2; Length 2704;  
Best Local Similarity 20.0%; Pred. No. 4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
DB 1632 CVAATATATTC 1641

RESULT 12  
F84312  
hypothetical protein Vng1598h [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: F84312  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laeky, S  
; Leithauser, B.; Keller, K.; Cruz, R.; Danison, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebdardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: F84312  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-71 <STO>  
A:Cross-references: UNIPROT:Q9HPL; UNIPARC:UPI0000063920; GB:AE004437; NID:g10581076; P  
C:Genetics:  
A:Gene: VNG1598H

Query Match 50.0%; Score 16; DB 2; Length 71;  
Best Local Similarity 20.0%; Pred. No. 4.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
DB 30 CSRSVSTSTYC 39

RESULT 13  
E70531  
hypothetical protein RV2706c - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: E70531  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S  
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentile, S.; Hamlin, N.; Holtrold, S  
; Raftery, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: E70531  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-85 <COL>  
A:Cross-references: UNIPROT:O07207; UNIPARC:UPI00000C14DC; GB:Z6072; GB:AL123456; NID:g  
C:Genetics:  
A:Experimental source: strain H37RV  
A:Gene: RV2706c

Query Match 50.0%; Score 16; DB 2; Length 85;  
Best Local Similarity 20.0%; Pred. No. 5;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
DB 24 GSATVAVAAC 33

RESULT 14  
S51479  
drought-induced protein D121 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 15-Jul-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S51479; S43176

R.Gosti, F.; Bertauche, N.; Vartanian, N.; Giraudat, J.

Mol. Gen. Genet. 246, 10-18, 1995

A:Title: Abscisic acid-dependent and -independent regulation of gene expression by proga

A:Reference number: S51478; MID:95124230; PMID:7823904

A:Accession: S51479

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-104 <GOS>

A:Cross-references: UNIPROT:Q39084; UNIPARC:UPI000000BF21; EMBL:X78585; NID:9469111; PID

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994

C:Genetics:

A:Gene: D121

C:Superfamily: late embryogenesis-abundant protein leas

Query Match

Best Local Similarity 50.0%; Score 16; DB 2; Length 104;  
Pred. No. 5.4;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 14 CSASGSLSC 23

RESULT 15

A71249

hypothetical protein PH0248 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 12-Jul-2004

C:Accession: A71249

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A:Reference number: A71000; MID:98344157; PMID:9679194

A:Accession: A71249

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-105 <KAW>

A:Cross-references: UNIPROT:Q57986; UNIPARC:UPI0000062D66; GB:AP000001; NID:93236128; PI

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH0248

Query Match

Best Local Similarity 50.0%; Score 16; DB 2; Length 105;  
Pred. No. 5.4;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 35 CFTASASLAC 44

RESULT 16

F72549

hypothetical protein APE1682 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C:Accession: F72549

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MID:99310339; PMID:10382966

A:Accession: F72549

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-108 <KAW>

A:Cross-references: UNIPROT:Q9YB85; UNIPARC:UPI000005E048; DDBJ:AP0000062; NID:95105244;

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1682

C:Superfamily: Aeropyrum pernix hypothetical protein APE1682

Query Match

Best Local Similarity 50.0%; Score 16; DB 2; Length 108;  
Pred. No. 5.5;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 8 CFSASKAAC 17

RESULT 17

A23473

chymotrypsin-like proteinase (EC 3.4.21.-) - pig (tentative sequence) (fragments)

N:Alternate names: pancreatic elastase II [misidentification]

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 09-Jul-2004

C:Accession: A23473

R:Vered, M.; Gertler, A.; Burshtein, Y.

Int. J. Pept. Protein Res. 27, 183-190, 1986

A:Reference number: A23473; MID:86194534; PMID:3634756

A:Accession: A23473

A:Molecule type: protein

A:Residues: 1-126 <VER>

A:Cross-references: UNIPROT:Q7M325; UNIPARC:UPI0000175C02

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

Query Match

Best Local Similarity 50.0%; Score 16; DB 2; Length 126;  
Pred. No. 5.8;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 87 CAGASGASSC 96

RESULT 18

T49498

hypothetical protein B14D6.500 [imported] - Neurospora crassa

C:Species: Neurospora crassa

C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000

C:Accession: T49498

R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022

A:Accession: T49498

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-129 <SCH>

A:Cross-references: UNIPARC:UPI0000179E1D; EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.500

A:Experimental source: BAC clone B14D6; strain OR74A

C:Genetics:

A:Gene: NCSP:B14D6.500

A:Map position: 6

C:Superfamily: Neurospora crassa hypothetical protein B14D6.500

Query Match

Best Local Similarity 50.0%; Score 16; DB 2; Length 129;  
Pred. No. 5.8;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 90 CSARASISRC 99

RESULT 19

A96746

hypothetical protein T9N14.5 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: A96746

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.: Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Schmitt, P.; Southwick, A.M.; Sun, H.; Tallon,  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: A96746  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-133 <STO>  
 A:Cross-references: UNIPROT:Q9C7T1; UNIPARC:UPI000004A50D; GB:AE005173; NID:G10645374; F  
 C:Genetics:  
 A:Gene: T9N14.5  
 A:Map position: 1

Query Match 50.0%; Score 16; DB 2; Length 133;  
 Best Local Similarity 20.0%; Pred. No. 5.9;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
 DB 93 CYSSTSSVSC 102

RESULT 20  
 S09762  
 hypothetical protein TRL3 precursor - human cytomegalovirus (strain AD169)  
 C:Species: human cytomegalovirus, human herpesvirus 5  
 A:Name: host Homo sapiens (man)  
 C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
 C:Accession: S09762  
 R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;  
 M.; Barrall, B.G.  
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990  
 A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus  
 A:Reference number: S09749; MUID:90269039; PMID:2161319  
 A:Accession: S09762  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-147 <CHE>  
 A:Cross-references: UNIPROT:P16811; UNIPARC:UPI000012D86F; EMBL:X17403; NID:G59591; PIDN  
 A:Note: this sequence was submitted to the EMBL Data Library, December 1989  
 C:Superfamily: human cytomegalovirus hypothetical protein TRL3  
 C:Keywords: glycoprotein  
 F:1-35/Domains: signal sequence #status predicted <SIG>  
 F:26-147/Product: hypothetical protein TRL3 #status predicted <MAT>  
 F:21,35,54,62,99,110,125,133,144/Binding site: carbohydrate (asn) (covalent) #status pre

Query Match 50.0%; Score 16; DB 2; Length 147;  
 Best Local Similarity 20.0%; Pred. No. 6.1;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
 DB 97 CTNTTTVTC 106

RESULT 21  
 PNO103  
 hypothetical 17k protein - barley stripe mosaic virus  
 C:Species: barley stripe mosaic virus, BSMV  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C:Accession: PNO103  
 R:Kozlov, Y.V.; Afanasiev, B.N.; Rupakov, V.V.; Golova, Y.B.; Kulaeva, O.I.; Dolja, V.V.  
 Mol. Biol. (Mosk.) 23, 1080-1090, 1989  
 A:Title: The complete nucleotide sequence of barley stripe mosaic virus RNA 3 and its va  
 A:Reference number: PNO102; MUID:90066400; PMID:2586501  
 A:Accession: PNO103  
 A:Molecule type: genomic RNA  
 A:Residues: 1-153 <KOZ>

A:Cross-references: UNIPROT:Q07118; UNIPARC:UPI00000F58A5  
 C:Genetics:  
 A:Map position: segment 3(III)

Query Match 50.0%; Score 16; DB 2; Length 153;  
 Best Local Similarity 20.0%; Pred. No. 6.2;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
 DB 10 CGTSTSTVC 19

RESULT 22  
 T28088  
 hypothetical protein ZK899.1 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T28088  
 R:Kershaw, J.  
 submitted to the EMBL Data Library, September 1994  
 A:Reference number: Z20468  
 A:Accession: T28088  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-161 <WIL>  
 A:Cross-references: UNIPROT:Q23659; UNIPARC:UPI000007B9F8; EMBL:Z37140; PIDN:CAA85496.1,  
 A:Experimental source: clone ZK899  
 C:Genetics:  
 A:Gene: CESP:ZK899.1  
 A:Map position: X 89/3; 144/2  
 A:Introns: 46/3;  
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein ZK899.1

Query Match 50.0%; Score 16; DB 2; Length 161;  
 Best Local Similarity 20.0%; Pred. No. 6.3;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
 DB 152 CSTSTSTVC 161

RESULT 23  
 J01252  
 hypothetical 16.7k protein - chrysanthemum virus B  
 C:Species: chrysanthemum virus B  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C:Accession: J01252  
 R:Levey, K.; Zayriev, S.  
 J. Gen. Virol. 72, 2333-2337, 1991  
 A:Title: Nucleotide sequence and gene organization of the 3'-terminal region of chrysant  
 A:Reference number: J01246; MUID:92013948; PMID:1919520  
 A:Accession: J01252  
 A:Molecule type: genomic RNA  
 A:Residues: 1-164 <LRV>  
 A:Cross-references: UNIPROT:Q7LZW7; UNIPARC:UPI0000179B53

Query Match 50.0%; Score 16; DB 2; Length 164;  
 Best Local Similarity 20.0%; Pred. No. 6.3;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
 DB 80 CSLSTSSASC 89

RESULT 24  
 T34277  
 hypothetical protein F46H5.2 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T34277

R:Nhan, M.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid F46H5.  
A:Reference number: Z21498

A:Accession: T34277  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-229 <SHA>  
A:Cross-references: UNIPROT:Q20490; UNIPARC:UPI0000179CAC; EMBL:U41543; PIDD:AA837021.1;  
A:Experimental source: strain Bristol NZ; clone F46H5  
C:Genetics:  
A:Gene: CESP:F46H5.2  
A:Map position: X  
A:Introns: 16/1; 85/2; 151/2

Query Match 50.0%; Score 16; DB 2; Length 229;  
Best Local Similarity 20.0%; Pred. No. 7.1;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 23 CEASASNSC 32

RESULT 25  
pectate lyase (EC 4.2.2.2) - fungus (Fusarium solani)  
C:Species: Fusarium solani  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A45724

R:Gonzalez-Candelas, L.; Kolattukudy, P.E.  
J. Bacteriol. 174, 6343-6349, 1992  
A:Title: Isolation and analysis of a novel inducible pectate lyase gene from the phytote  
A:Reference number: A45724; MUID:93015682; PMID:1400187  
A:Accession: A45724  
A:Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-242 <CON>  
A:Cross-references: UNIPROT:Q04701; UNIPARC:UPI000006C8B5; GB:M94691; NID:g168155; PIDD:  
A:Experimental source: isolate 78  
A:Note: sequence extracted from NCBI backbone (NCBIN:115473, NCBI:P115474)  
C:Keywords: carbon-oxygen lyase

Query Match 50.0%; Score 16; DB 2; Length 242;  
Best Local Similarity 20.0%; Pred. No. 7.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 233 CTVGTSTSC 242

RESULT 26  
chymotrypsin B - Atlantic cod (fragments)  
C:Species: Gadus morhua (Atlantic cod)  
C:Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 09-Jul-2004  
C:Accession: S72219  
R:Leith-Larsen, R.; Asgeirsson, B.; Thorolfsson, M.; Norregaard-Madsen, M.; Hojrup, P.  
Biochim. Biophys. Acta 1297, 49-56, 1996  
A:Title: Structure of chymotrypsin variant B from Atlantic cod, Gadus morhua.  
A:Reference number: S72219; MUID:96439045; PMID:8841380  
A:Accession: S72219  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-14;15-244 <LET>  
A:Cross-references: UNIPROT:Q9PM06; UNIPARC:UPI0000175C00; UNIPARC:UPI0000175C01  
C:Superfamily: trypsin; trypsin homology  
F:15-237/Domain: trypsin homology <TRY>

Query Match 50.0%; Score 16; DB 2; Length 244;  
Best Local Similarity 20.0%; Pred. No. 7.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 181 CAGAGATSC 190

RESULT 27  
A55035  
cysteine-rich protein CRP1 - earthworm (Enchytraeus buchholzi)  
C:Species: Enchytraeus buchholzi  
C:Date: 14-Nov-1994 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: A55035; S45034  
R:Willuhn, J.; Schmitt-Wrede, H.P.; Greven, H.; Wunderlich, F.  
J. Biol. Chem. 269, 24688-24691, 1994  
A:Title: cDNA cloning of a cadmium-inducible mRNA encoding a novel cysteine-rich, non-mc  
A:Reference number: A55035; MUID:95014230; PMID:7929141  
A:Accession: A55035  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-251 <ML>

A:Cross-references: UNIPROT:Q24774; UNIPARC:UPI000007D243; EMBL:X79344; NID:g488802; PIDD:  
C:Superfamily: ultra-high-sulfur keratin

Query Match 50.0%; Score 16; DB 2; Length 251;  
Best Local Similarity 20.0%; Pred. No. 7.4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 28 CGTSTTASC 37

RESULT 28  
J01724  
EI membrane glycoprotein precursor - canine coronavirus (strain Insavc-1)  
N:Alternate names: matrix glycoprotein  
C:Species: canine coronavirus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 09-Jul-2004  
C:Accession: J01724  
R:Horzburgh, B.C.; Brierley, I.; Brown, T.D.K.  
J. Gen. Virol. 73, 2849-2862, 1992  
A:Title: Analysis of a 9.6 kb sequence from the 3' end of canine coronavirus genomic RNA  
A:Reference number: P00481; MUID:93057357; PMID:1431811  
A:Accession: J01724  
A:Molecule type: genomic RNA  
A:Residues: 1-262 <HOR>  
A:Cross-references: UNIPROT:P36299; UNIPARC:UPI0000138986; DDBJ:D13096; NID:g406193; PIDD:  
C:Genetics:  
A:Gene: M

C:Superfamily: coronavirus EI membrane glycoprotein  
C:Keywords: glycoprotein; matrix protein; transmembrane protein  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-262/Product: EI membrane glycoprotein #status predicted <MAT>  
F:56-72/Domain: transmembrane #status predicted <TM1>  
F:85-101/Domain: transmembrane #status predicted <TM2>  
F:115-134/Domain: transmembrane #status predicted <TM3>  
F:32,55/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:33,57/Binding site: carbohydrate (Ser) (covalent) #status predicted  
F:44/Binding site: carbohydrate (Thr) (covalent) #status predicted

Query Match 50.0%; Score 16; DB 1; Length 262;  
Best Local Similarity 20.0%; Pred. No. 7.5;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 21 CAMTSSTSC 30

RESULT 29  
S47537  
chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod  
C:Species: Gadus morhua (Atlantic cod)

C.Date: 26-Dec-1994 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004  
C.Accession: S47537, S43163  
R.Gundmundsdottir, A.; Oskarsson, S.; Eakin, A.E.; Craik, C.S.; Bjarnason, J.B.  
Biochim. Biophys. Acta 1219, 211-214, 1994  
A.Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.  
A.Reference number: S47537, MUID:94368860, PMID:8086467  
A.Accession: S47537  
A.Molecule type: mRNA  
A.Residues: 1-263 <GUD>  
A.Cross-references: UNIPROT:P47796, UNIPARC:UPI0000128664, EMBL:X78490, NID:9468750, PID  
C:Superfamily: trypsin, trypsin homology  
C:Keywords: hydrolase; protein digestion; serine proteinase  
F.1-18/Domain: signal sequence #status predicted <SIG>  
F.19-263/Product: chymotrypsin #status predicted <MAT>  
F.34-256/Domain: trypsin homology <TRY>  
F.75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 50.0%; Score 16; DB 2; Length 263;  
Best Local Similarity 20.0%; Pred. No. 7.5;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
DB 200 CAGAGASSC 209

RESULT 30  
A5856  
Probable elongation factor yelp [imported] - Escherichia coli (strain O157:H7, substrain  
C:Species: Escherichia coli  
C.Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 05-Oct-2004  
C.Accession: A5856  
R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; DiMantata, E.; Potamoustis, K.; Apodaca,  
Natter 409, 529-533, 2001  
A.Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A.Reference number: A58480, MUID:21074935, PMID:11206551  
A.Accession: A5856  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-275 <STO>  
A.Cross-references: UNIPARC:UPI00000D0288, GB:AE005174, NID:912516497, PIDN:AAG57309.1;  
A.Experimental source: strain O157:H7, substrain EDL933  
C.Genetics:  
A.Gene: yelp  
C:Superfamily: translation elongation factor P

Query Match 50.0%; Score 16; DB 2; Length 275;  
Best Local Similarity 20.0%; Pred. No. 7.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
DB 54 CTSKSANTTC 63

RESULT 31  
B64986  
Hypothetical 30.9 kD protein in frub 5' region - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C.Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 05-Oct-2004  
C.Accession: B64986  
R.Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A.Title: The complete genome sequence of Escherichia coli K-12.  
A.Reference number: A64720, MUID:97426617, PMID:9278503  
A.Accession: B64986  
A.Status: preliminary; nucleic acid sequence not shown; translation not shown  
A.Molecule type: DNA  
A.Residues: 1-275 <BLAT>  
A.Cross-references: UNIPARC:UPI0000167DC1, GB:AE000306, GB:U000096, NID:91788489, PIDN:AA  
A.Experimental source: strain K-12, substrain MG1655

C.Genetics:  
A.Gene: yelp  
C:Superfamily: translation elongation factor P

Query Match 50.0%; Score 16; DB 2; Length 275;  
Best Local Similarity 20.0%; Pred. No. 7.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
DB 54 CTSKSANTTC 63

RESULT 32  
G91011  
Probable elongation factor [imported] - Escherichia coli (strain O157:H7, substrain RMD  
C:Species: Escherichia coli  
C.Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 05-Oct-2004  
C.Accession: G91011  
R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A.Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
A.Reference number: A99629, MUID:21156231, PMID:11258796  
A.Accession: G91011  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-275 <HAY>  
A.Cross-references: UNIPARC:UPI00000D0288, GB:BA000007, PIDN:BA836486.1, PID:913362532;  
A.Experimental source: strain O157:H7, substrain RMD 0509952  
C.Genetics:  
A.Gene: RC93063  
C:Superfamily: translation elongation factor P

Query Match 50.0%; Score 16; DB 2; Length 275;  
Best Local Similarity 20.0%; Pred. No. 7.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
DB 54 CTSKSANTTC 63

RESULT 33  
T23682  
Hypothetical protein M02G9.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C.Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C.Accession: T23682  
R.Mathews, L.  
Submitted to the EMBL Data Library, November 1996  
A.Reference number: Z19781  
A.Accession: T23682  
A.Status: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: DNA  
A.Residues: 1-294 <MIL>  
A.Cross-references: UNIPROT:O9XUS0, UNIPARC:UPI0000076A8F, EMBL:Z81573, PIDN:CAB04626.1;  
A.Experimental source: clone M02G9  
C.Genetics:  
A.Gene: CESP:M02G9.3  
A.Map position: 2  
A.Introns: 20/3; 76/2; 182/3; 223/3

Query Match 50.0%; Score 16; DB 2; Length 294;  
Best Local Similarity 20.0%; Pred. No. 7.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
DB 43 CSSSSSNSTC 52

RESULT 34



C71498  
probable Yop translocation R - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C/Accession: Chlamydia trachomatis  
C/Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004  
C/Accession: C71498  
R/Stephen, R.S.; Kaiman, S.; Lammell, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
A/Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
A/Reference number: A71570; UID:99000809; PMID:9784136  
A/Accession: C71498  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-306 <ARN>  
A/Cross-references: UNIPROT:084566; UNIPARC:UPI000003360; GB:AE001327; GB:AE001273; NID:94056491; P/Genetics:  
A/Experimental source: serotype D, strain UW-3/Cx  
A/Genes: yecR

Query Match 50.0%; Score 16; DB 2; Length 306;  
Best Local Similarity 20.0%; Pred. No. 7.9;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 24 CASACGSSC 33

RESULT 35  
T00500  
probable elicitor response element-binding protein WRKY3 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 31-Dec-2004  
C/Accession: T00500; C84623  
R/Rounsley, S.D.; Jin, X.; Kechum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul, submitted to the EMBL Data Library, November 1997  
A/Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.  
A/Reference number: Z14164  
A/Accession: T00500  
A/Status: translated from GB/EMBL/DBDUT  
A/Molecule type: DNA  
A/Residues: 1-317 <RNU>  
A/Cross-references: UNIPROT:Q22176; UNIPARC:UPI0000138F7D; EMBL:AC002391; NID:92642427;  
A/Experimental source: cultivar Columbia  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Neuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; UID:20083487; PMID:10617197  
A/Accession: C84623  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-317 <STO>  
A/Cross-references: UNIPARC:UPI0000138F7D; GB:AE002093; NID:92642432; PIDN:AAB87100.1; G/Genetics:  
A/Genes: WRKY3; ATSP:T20D16.5; At2g23320  
A/Map position: 2 260/2  
A/Intons: 218/2; DNA-binding protein WRKY3

Query Match 50.0%; Score 16; DB 2; Length 317;  
Best Local Similarity 20.0%; Pred. No. 8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 203 CASASSGRC 212

RESULT 36  
A84792  
hypothetical protein At2g37380 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: A84792  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Neuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; UID:20083487; PMID:10617197  
A/Accession: A84792  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-321 <STO>  
A/Cross-references: UNIPROT:Q9ZUS8; UNIPARC:UPI00000A2519; GB:AE002093; NID:94056491; P/Genetics:  
A/Genes: At2g37380  
A/Map position: 2  
C/Superfamily: Arabidopsis thaliana hypothetical protein At2g39370

Query Match 50.0%; Score 16; DB 2; Length 321;  
Best Local Similarity 20.0%; Pred. No. 8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 306 CSSRTSVSTC 315

RESULT 37  
J0202  
paired-box-containing Pax4 protein - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C/Accession: J0202  
R/Tokuyama, Y.; Yagui, K.; Sakurai, K.; Hashimoto, N.; Saito, Y.; Kanatsuka, A. Biochem. Biophys. Res. Commun. 248, 153-156, 1998  
A/Title: Molecular cloning of rat Pax4: identification of four isoforms in rat insulinoma  
A/Reference number: J0202; UID:98340866; PMID:9675102  
A/Accession: J0202  
A/Molecule type: mRNA  
A/Residues: 1-349 <TOX>  
A/Cross-references: UNIPROT:088436; UNIPARC:UPI000013136B; GB:AF053100; NID:93192973; P/Genetics:  
C/Superfamily: paired box transcription factor Pax-4; homeobox homology; paired box homology  
C/Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F/5-129/Domain: paired box homology <PBH>  
F/171-227/Domain: homeobox homology <HOX>

Query Match 50.0%; Score 16; DB 2; Length 349;  
Best Local Similarity 20.0%; Pred. No. 8.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 290 CSSDTTSGAC 299

RESULT 38  
T04007  
hypothetical protein T5L19.160 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004  
C/Accession: T04007  
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X submitted to the Protein Sequence Database, March 1999  
A/Reference number: Z15184  
A/Accession: T04007  
A/Molecule type: DNA  
A/Residues: 1-380 <BEV>  
A/Cross-references: UNIPROT:Q9T0G1; UNIPARC:UPI00000A0088; EMBL:AL049481  
A/Experimental source: cultivar Columbia; BAC clone T5L19  
C/Genetics:  
A/Map position: 4  
A/Intons: 38/3; 83/3, 92/3, 123/1, 137/3, 170/3, 190/1, 207/3, 241/3, 263/3, 301/2, 358/3  
A/Note: T5L19.160

Query Match 50.0%; Score 16; DB 2; Length 380;  
Best Local Similarity 20.0%; Pred. No. 8.5;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 6 CSTASSGGC 15

RESULT 39  
T26077  
hypochemical protein W02A2.3 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T26077  
R:Ainscough, R.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z20148  
A:Accession: T26077  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-397 <MIL>  
A:Cross-references: UNIPROT:Q9XUB6; UNIPARC:UPI0000076467; EMBL:Z82286; PIDD: CAB05306.1;  
C:Experimental source: clone W02A2  
C:Genetics:  
A:Gene: CESP.W02A2.3  
A:Map position: 4  
A:Introns: 15/1; 67/2; 101/3; 259/1; 383/3  
C:Superfamily: gliadin

Query Match 50.0%; Score 16; DB 2; Length 397;  
Best Local Similarity 20.0%; Pred. No. 8.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 10 CATLAALASC 19

RESULT 40  
S41514  
RAD52 protein homolog - chicken  
C:Species: *Gallus gallus* (chicken)  
C:Date: 25-Dec-1994 #sequence\_revision 10-Nov-1995 #text\_change 31-Dec-2004  
C:Accession: S41514  
R:Bezubova, O.Y.; Schmidt, H.; Ostermann, K.; Heyer, W.D.; Buerstedde, J.M.  
Nucleic Acids Res. 21, 5945-5949, 1993  
A:Title: Identification of a chicken RAD52 homologue suggests conservation of the RAD52  
A:Reference number: S41514; MUID:94119695; PMID:8290357  
A:Accession: S41514  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-422 <BEZ>  
A:Cross-references: UNIPROT:P39022; UNIPARC:UPI0000133034  
C:Superfamily: human nuclear cap-binding protein

Query Match 50.0%; Score 16; DB 2; Length 422;  
Best Local Similarity 20.0%; Pred. No. 8.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 17 CTSTNSVAC 26

RESULT 41  
C81039  
lipopolysaccharide biosynthesis protein, probable MMB1818 [imported] - *Neisseria meningi*  
C:Species: *Neisseria meningitidis*  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: C81039  
R:Retelien, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Bisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
Zi, H.; Qin, H.; Vamathavan, J.; Gill, J.; Scarlato, V.; Masiarant, V.; Pizza, M.  
Science 287, 1809-1815, 2000  
A:Authors: Grand, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.  
A:Reference number: A81000; MUID:20175755; PMID:10710307  
A:Accession: C81039  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-473 <TET>  
A:Cross-references: UNIPROT:Q9JXZ7; UNIPARC:UPI0000030867; GB:AE002531; GB:AE002098; NID  
A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NMB1818

Query Match 50.0%; Score 16; DB 2; Length 473;  
Best Local Similarity 20.0%; Pred. No. 9.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 421 CTTSSAAYTC 430

RESULT 42  
C81984  
Probable lipopolysaccharide biosynthesis translocase NMA0643 [imported] - *Neisseria men*  
C:Species: *Neisseria meningitidis*  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C:Accession: C81984  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.  
A:Reference number: A81775; MUID:20222556; PMID:10761919  
A:Accession: C81984  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-473 <PAR>  
A:Cross-references: UNIPROT:Q9JTW9; UNIPARC:UPI0000030868; GB:AL162753; GB:AL157959; NID  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: NMA0641; NMA0643

Query Match 50.0%; Score 16; DB 2; Length 473;  
Best Local Similarity 20.0%; Pred. No. 9.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 421 CLASSAAYTC 430

RESULT 43  
T36342  
probable glutamate decarboxylase - *Streptomyces coelicolor*  
C:Species: *Streptomyces coelicolor*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T36342  
R:Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, May 1999  
A:Reference number: Z21575  
A:Accession: T36342  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-475 <OLI>  
A:Cross-references: UNIPROT:Q9XGJ5; UNIPARC:UPI00000DB08A; EMBL:AL049841; PIDD: CAB42769.  
C:Genetics:  
A:Gene: gad; SCOEDB:SCB9.23  
C:Superfamily: *Escherichia coli* glutamate decarboxylase

Query Match 50.0%; Score 16; DB 2; Length 475;

Best Local Similarity 20.0%; Pred. No. 9.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 131 CTTGSSSAC 140

RESULT 44

B66460 hypothetical protein F14M2.19 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: B66460

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

anssen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Malt, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A66141; MUID:21016719; PMID:11130712

A:Accession: B66460

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-486 <STO>

A:Cross-references: UNIPROT:Q9LQ21; UNIPARC:UPI00000A7F31; GB:AE005172; NID:G9665100; PI

C:Genetics:

A:Map position: 1

Query Match 50.0%; Score 16; DB 2; Length 486;

Best Local Similarity 20.0%; Pred. No. 9.3;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 48 CUSSTGATC 57

RESULT 45

A55207 glycerol-3-phosphate dehydrogenase glpd (EC 1.1.1.-) - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa

C:Date: 05-May-1995 #sequence\_revision 05-May-1995 #text\_change 09-Jul-2004

C:Accession: A55207

R:Schweizer, H.P.; Po, C.

J. Bacteriol. 176, 2184-2193, 1994

A:Title: Cloning and nucleotide sequence of the glpd gene encoding sn-glycerol-3-phospha

A:Reference number: A55207; MUID:94209216; PMID:8157588

A:Accession: A55207

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-510 <SCH>

A:Cross-references: UNIPROT:P52111; UNIPARC:UPI000016FCAD; GB:L06231; NID:G450377; PIDN:

C:Genetics:

A:Gene: glpd

C:Superfamily: glycerol-3-phosphate dehydrogenase (aerobic)

C:Keywords: oxidoreductase

Query Match 50.0%; Score 16; DB 2; Length 510;

Best Local Similarity 20.0%; Pred. No. 9.4;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 194 CTVSAAATAC 203

RESULT 46

G31277

guinate transport protein - Neurospora crassa (tentative sequence)

N:Alternate names: guinate transporter

C:Species: Neurospora crassa

C:Date: 26-Apr-1989 #sequence\_revision 26-Apr-1989 #text\_change 09-Jul-2004

C:Accession: S04254; G31277

R:Geever, R.F.; Hulet, L.; Baum, J.A.; Tyler, B.M.; Patel, V.B.; Rutledge, B.J.; Case, P.

J. Mol. Biol. 207, 15-34, 1989

A:Title: DNA sequence, organization and regulation of the ga gene cluster of Neurospora

A:Reference number: S04250; MUID:89293848; PMID:2525625

A:Accession: S04254

A:Molecule type: DNA

A:Residues: 1-537 <GE2>

A:Cross-references: UNIPROT:P1636; UNIPARC:UPI0000132E4B; EMBL:X14603; NID:G3060; PIDN:

C:Genetics:

A:Gene: ga-Y

C:Superfamily: maltose transport protein MAL61

C:Keywords: transmembrane protein

F:162-42/Domain: transmembrane #status predicted <TM01>

F:167-87/Domain: transmembrane #status predicted <TM02>

F:199-119/Domain: transmembrane #status predicted <TM03>

F:132-155/Domain: transmembrane #status predicted <TM04>

F:161-181/Domain: transmembrane #status predicted <TM05>

F:195-215/Domain: transmembrane #status predicted <TM06>

F:286-306/Domain: transmembrane #status predicted <TM07>

F:324-344/Domain: transmembrane #status predicted <TM08>

F:356-376/Domain: transmembrane #status predicted <TM09>

F:390-410/Domain: transmembrane #status predicted <TM10>

F:427-447/Domain: transmembrane #status predicted <TM11>

F:459-479/Domain: transmembrane #status predicted <TM12>

Query Match 50.0%; Score 16; DB 2; Length 537;

Best Local Similarity 20.0%; Pred. No. 9.6;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 24 CAATSPASC 33

RESULT 47

B38418 jockey protein 1 - fruit fly (Drosophila funebris)

C:Species: Drosophila funebris

C:Date: 28-Jun-1991 #sequence\_revision 28-Jun-1991 #text\_change 21-Feb-1997

C:Accession: B38418

R:Itizokhi, L.U.; Mazo, A.M.

Proc. Natl. Acad. Sci. U.S.A. 87, 9216-9220, 1990

A:Title: Evidence for horizontal transmission of the mobile element jockey between dista

A:Reference number: A38418; MUID:91067678; PMID:1701254

A:Accession: B38418

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-580 <MT2>

A:Cross-references: UNIPARC:UPI000017BE6C; GB:M38437

C:Genetics:

A:Gene: FlyBase:Dfun/jockey

A:Cross-references: FlyBase:FBgn012310

Query Match 50.0%; Score 16; DB 2; Length 580;

Best Local Similarity 20.0%; Pred. No. 9.9;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

DB 78 CTVTTSQASC 87

RESULT 48

T22518

hypothetical protein FS2H3.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T22518

R.Gardner, A.  
submitted to the EMBL Data Library, October 1995  
A:Accession number: Z19575  
A:Accession: T22518  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-638 <WIL>  
A:Cross-references: UNIPROT:Q20680; UNIPARC:UPI000013C10D; EMBL:Z66512; PIDD:CAA91322.1;  
A:Experimental source: clone F52H3  
C:Genetics:  
A:Gene: CESP:F52H3.2  
A:Map position: 2  
A:Introns: 67/1; 157/2; 285/2; 420/1; 566/3  
C:Superfamily: gida protein

Query Match 50.0%; Score 16; DB 2; Length 638;  
Best Local Similarity 20.0%; Pred. No. 10;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 13  
DB 29 CESAARARC 38

RESULT 49  
C81345  
hypothetical protein Cj0736 [imported] - Campylobacter jejuni (strain NCTC 11168)  
C:Species: Campylobacter jejuni  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: C81345  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillif  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell  
Nature 403, 665-668, 2000  
A>Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: A81250; MUID:20150912; PMID:10688204  
A:Accession: C81345  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-642 <PAR>  
A:Cross-references: UNIPROT:Q9PPH1; UNIPARC:UPI00000C1D12; GB:AL139076; GB:AL111168; NID  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj0736

Query Match 50.0%; Score 16; DB 2; Length 642;  
Best Local Similarity 20.0%; Pred. No. 10;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 13  
DB 375 CSSAAVSSC 384

RESULT 50  
T37581  
probable serine-threonine-protein kinase - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T37581  
R:Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, April 1996  
A:Reference number: Z21726  
A:Accession: T37581  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-648 <BAD>  
A:Cross-references: UNIPARC:UPI000017B219; EMBL:Z70721; PIDD:CAA94704.1; GSPDB:GN00066;  
C:Genetics:  
A:Gene: SPDB:SPAC12B10.14C  
A:Map position: 1

Query Match 50.0%; Score 16; DB 2; Length 648;

Best Local Similarity 20.0%; Pred. No. 10;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 13  
DB 21 COSNASSAC 30

Search completed: January 4, 2006, 16:10:17  
Job time : 25.5565 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 104.487 Seconds

(without alignments)  
108.037 Million cell updates/sec

Title: US-09-932-322-4

Perfect score: 32

Sequence: 1 XXXXXXXXXXXXXXX 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : UniProt\_05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	56.2	213	2	Q96RM3 HUMAN
2	18	56.2	263	2	Q586F4 YTRYP
3	18	56.2	269	2	Q7PR07 ANOGA
4	18	56.2	281	2	Q5TX10 ANOGA
5	18	56.2	313	1	YMS8 YEAST
6	18	56.2	357	2	Q6ZDR4 ORYSA
7	18	56.2	362	2	Q7QJL0 ANOGA
8	18	56.2	480	2	Q9RV87 DEIRA
9	18	56.2	653	1	TMCC1 HUMAN
10	18	56.2	1005	2	Q5H317 XANOR
11	18	56.2	1061	2	Q4QCK7 LEIMA
12	17	53.1	52	2	Q8LPP6 HORVU
13	17	53.1	122	2	Q5DCN3 SCHJA
14	17	53.1	128	2	Q96LJ4 HUMAN
15	17	53.1	139	2	Q6ZKM7 ORYSA
16	17	53.1	155	2	Q6H482 ORYSA
17	17	53.1	160	2	Q657M0 ORYSA
18	17	53.1	160	2	Q4T4H1 TETNG
19	17	53.1	219	2	Q825E8 STRAM
20	17	53.1	220	2	Q66152 GVRIRU
21	17	53.1	256	2	Q6YTS2 ORYSA
22	17	53.1	260	2	Q850Z3 ORYSA
23	17	53.1	304	2	Q5TVE4 ANOGA
24	17	53.1	344	2	Q582B8 YTRYP
25	17	53.1	361	2	Q57220 YTRYP
26	17	53.1	364	2	Q4Y877 PLABE
27	17	53.1	368	2	Q70RD3 GERBY
28	17	53.1	376	1	CBID AGRTS
29	17	53.1	388	2	Q7PY13 ANOGA
30	17	53.1	389	2	Q21081 CAENORHABDI
31	17	53.1	411	2	Q67UZ3 ORYSA

32	17	53.1	437	2	Q414Q8 GIBBE
33	17	53.1	471	2	Q9VMG7 DROME
34	17	53.1	475	2	Q7X7A4 ORYSA
35	17	53.1	491	2	Q4P4C7 USHMA
36	17	53.1	520	2	Q4XZV5 PLABE
37	17	53.1	556	2	Q803D5 XANOR
38	17	53.1	568	2	Q5GXM1 XANOR
39	17	53.1	626	2	Q4N2Y7 THEPA
40	17	53.1	630	2	Q69YV8 ORYSA
41	17	53.1	634	1	HSP71 LEIMA
42	17	53.1	635	2	Q4Q747 LEIMA
43	17	53.1	650	2	Q4Q740 LEIMA
44	17	53.1	662	2	Q4Q744 LEIMA
45	17	53.1	662	2	Q4Q745 LEIMA
46	17	53.1	691	2	Q60E18 ORYSA
47	17	53.1	695	2	Q819P8 YTRYP
48	17	53.1	705	2	Q516X2 CHLAB
49	17	53.1	734	2	Q9NKE3 DROME
50	17	53.1	784	2	Q4YNR2 PLABE
51	17	53.1	843	2	Q80GV2 HEPAT
52	17	53.1	861	2	Q55NL5 CRYNE
53	17	53.1	943	1	BLI4 CABE
54	17	53.1	1089	2	Q4Q9C2 LEIMA
55	17	53.1	1103	2	Q4PDJ1 USHMA
56	17	53.1	1200	2	Q4RF81 TETNG
57	17	53.1	1347	2	Q6CC16 YARLI
58	17	53.1	1749	2	Q4V217 BURMA
59	17	53.1	1835	2	Q4S914 TETNG
60	17	53.1	2704	2	G168 PARR
61	17	53.1	2717	2	Q94710 PARTE
62	17	53.1	2721	2	Q76973 PARTE
63	17	53.1	2729	2	Q6P0K6 PARTE
64	16	50.0	68	2	Q6JL92 NEI6O
65	16	50.0	71	2	Q9HPK1 HALBA
66	16	50.0	80	2	Q6IGD0 DROME
67	16	50.0	85	2	Q07207 MYCBO
68	16	50.0	85	2	Q7RY24 MYCBO
69	16	50.0	85	2	Q7NSB4 CHRVO
70	16	50.0	88	2	Q5Z713 ORYSA
71	16	50.0	98	2	Q5Z2E5 NOCHA
72	16	50.0	103	2	Q5YF63 YTRYP
73	16	50.0	104	2	Q39084 ARATH
74	16	50.0	105	2	Q57986 PYRHO
75	16	50.0	108	2	Q9YBBS ARABE
76	16	50.0	108	2	Q85429 YCALUD
77	16	50.0	110	2	Q6Y9M8 LIGOV
78	16	50.0	112	2	Q7PJV4 ANOGA
79	16	50.0	112	2	Q95M57 BOVIN
80	16	50.0	112	2	Q8V6M9 YVIRU
81	16	50.0	120	2	Q6ID11 ARATH
82	16	50.0	123	2	Q84WT9 ARATH
83	16	50.0	125	2	Q59Z22 CANAL
84	16	50.0	126	2	Q7M325 PIGE
85	16	50.0	126	2	Q6YUC3 ORYSA
86	16	50.0	126	2	Q8CLV7 YERPE
87	16	50.0	129	2	Q7RUV2 NEURC
88	16	50.0	133	2	Q6UNK8 BOMBO
89	16	50.0	133	2	Q8C7T1 ARATH
90	16	50.0	137	2	Q9HAI1 HUMAN
91	16	50.0	137	2	Q24987 GIALA
92	16	50.0	137	2	Q8WLI9 DROME
93	16	50.0	142	2	Q08497 HUMAN
94	16	50.0	147	1	IRI3 HGMVA
95	16	50.0	147	2	Q6RE16 SNCCA
96	16	50.0	149	2	Q4RZ65 TETNG
97	16	50.0	149	2	Q4S237 TETNG
98	16	50.0	151	2	Q5H3T3 XANOR
99	16	50.0	153	1	GAMB BSNV
100	16	50.0	157	2	Q8LI53 ORYSA
101	16	50.0	158	2	Q8CC93 MOUSE
102	16	50.0	161	2	Q236S9 CABEL
103	16	50.0	162	2	Q91G16 ADEP3
104	16	50.0	164	2	Q7LZW7 CVB

Q414Q8 GIBBERELLA	Q9VMG7 DROSOPHILA	Q7X7A4 ORYSA	Q4P4C7 USHLAGO MA	Q4XZV5 PLASMODIUM	Q803D5 BRACHYDANIO	Q5GXM1 XANTHOMONAS	Q4N2Y7 THELLERIA P	Q69YV8 ORYZA SATIV	Q4Q747 LEISHMANIA	Q4Q740 LEISHMANIA	Q4Q744 LEISHMANIA	Q4Q745 LEISHMANIA	Q60E18 ORYZA SATIV	Q819P8 CORBICULA F	Q516X2 CHLAMYDOPHIL	Q9NKE3 DROSOPHILA	Q4YNR2 PLASMODIUM	Q80GV2 HEPATOCYTE	Q55NL5 CRYPTOCOCCUS	P51559 CAENORHABDITI	Q4Q9C2 LEISHMANIA	Q4PDJ1 USHLAGO MA	Q4RF81 TETRAODON N	Q6CC16 YARROWIA LI	Q4V217 BURKHOLDERIA	Q4S914 TETRAODON N	P17053 PARAMETICIUM	Q94710 PARAMETICIUM	Q76973 PARAMETICIUM	Q6P0K6 PARAMETICIUM	Q6JL92 NEISSERIA G	Q9HPK1 HALOBACTERIUM	Q6IGD0 DROSOPHILA	Q07207 MYCOBACTERIUM	Q7RY24 MYCOBACTERIUM	Q7NSB4 CHROMOBACTERIUM	Q5Z713 ORYZA SATIV	Q5Z2E5 NOCARDIA FA	Q5YF63 ROCK BREAM	Q39084 ARABIDOPSIS	Q57986 PYROCOCCUS	Q9YBBS ACETOXYLUM P	Q85429 MYCOBACTERIUM	Q6Y9M8 LIGUSTRUM O	Q7PJV4 ANOPHELES G	Q95M57 BOS TAURUS	Q8V6M9 HALOVLIVUS H	Q6ID11 ARABIDOPSIS	Q84WT9 ARABIDOPSIS	Q59Z22 CANDIDA ALB	Q7M325 SUS SCROFA	Q6YUC3 ORYZA SATIV	Q8CLV7 YERSINIA PE	Q7RUV2 NEUROSPORA	Q6UNK8 BOMBYX MORI	Q8C7T1 ARABIDOPSIS	Q9HAI1 HOMO SAPIEN	Q24987 GLADIOLA LAM	Q8WLI9 DIOSPHILA	Q08497 HOMO SAPIEN	P16811 HOMO CYTOM	Q6RE16 RHODOCOCCUS	Q4RZ65 TETRAODON N	Q4S237 TETRAODON N	Q5H3T3 XANTHOMONAS	Q80873 BARLEY STRI	Q8LI53 ORYZA SATIV	Q8CC93 MUS MUSCULU	Q236S9 CAENORHABDITI	Q91G16 PORCINE ADE	Q7LZW7 CHRYSANTHEMUM
-------------------	-------------------	--------------	-------------------	-------------------	--------------------	--------------------	--------------------	--------------------	-------------------	-------------------	-------------------	-------------------	--------------------	--------------------	---------------------	-------------------	-------------------	-------------------	---------------------	----------------------	-------------------	-------------------	--------------------	--------------------	---------------------	--------------------	---------------------	---------------------	---------------------	---------------------	--------------------	----------------------	-------------------	----------------------	----------------------	------------------------	--------------------	--------------------	-------------------	--------------------	-------------------	---------------------	----------------------	--------------------	--------------------	-------------------	---------------------	--------------------	--------------------	--------------------	-------------------	--------------------	--------------------	-------------------	--------------------	--------------------	--------------------	---------------------	------------------	--------------------	-------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	--------------------	----------------------	--------------------	----------------------

105	16	50.0	166	2	061J45_DROME	061J45_drosophila	178	16	50.0	272	2	06S5G8_9CNTD	06S5g8_nematostell
106	16	50.0	167	2	024969_GIALA	024969_giardia lam	179	16	50.0	272	2	06K6J9_9CNTD	06K6j9_nematostell
107	16	50.0	170	2	09B1L7_GIALA	09b1l7_giardia lam	180	16	50.0	276	2	04W1J3_ASPFU	04w1j3_aspergillus
108	16	50.0	170	2	08XSZ7_RALSO	08xs27_ralstonia s	181	16	50.0	280	2	06ZRL6_HUMAN	06zrl6_homo sapien
109	16	50.0	171	2	06ZP60_HUMAN	06z60_homo sapien	182	16	50.0	280	2	06ZUB8_HUMAN	06zub8_homo sapien
110	16	50.0	171	2	06OZL5_CABER	06ozl5_caenorhabdi	183	16	50.0	281	2	07ZHQ7_9H1V1	07zhq7_human immun
111	16	50.0	171	2	08MOAO_CAEEL	08moa0_caenorhabdi	184	16	50.0	282	1	KR101_HUMAN	060331_homo sapien
112	16	50.0	171	2	058S81_PSEFU	05881_pseudomonas	185	16	50.0	284	2	08MYV6_DROME	08myv6_drosophila
113	16	50.0	172	2	0988Y7_RHIL0	0988y7_rhizobium l	186	16	50.0	285	2	08S902_SOYBN	08s902_glycine max
114	16	50.0	174	2	071D98_DROYA	071d98_drosophila	187	16	50.0	287	2	08WVJ7_9ASCI	08wv7_boletia vi
115	16	50.0	175	2	09VSQ7_DROME	09vsq7_drosophila	188	16	50.0	287	2	04X7N6_PLACH	04x7n6_plasmodium
116	16	50.0	175	2	04JMM7_9BACT	04jmm7_uncultured	189	16	50.0	287	2	05U140_DROME	05u140_drosophila
117	16	50.0	179	2	04MD46_9BURK	04md46_burkholderi	190	16	50.0	290	2	086P14_DROME	086p14_drosophila
118	16	50.0	181	2	06ZTR4_HUMAN	06ztr4_homo sapien	191	16	50.0	290	2	04FX06_LEIMA	04fx06_leishmania
119	16	50.0	188	2	04RZL3_TETNG	04rzl3_tetradodon n	192	16	50.0	290	2	06RHX7_LYCES	06rhx7_lycopersico
120	16	50.0	189	2	05GXC7_XANOR	05gxc7_xanthomonas	193	16	50.0	291	2	05B3M4_EMENT	05b3m4_aspergillus
121	16	50.0	191	2	07OOR6_GIALA	07oor6_giardia lam	194	16	50.0	294	2	0620N2_CABER	0620n2_caenorhabdi
122	16	50.0	192	2	080AAB_TOSV	080aa8_toscana vir	195	16	50.0	294	2	09XUS0_CABER	09xus0_caenorhabdi
123	16	50.0	203	2	06ZVU4_HUMAN	06zvu4_homo sapien	196	16	50.0	294	2	04RJG2_TETNG	04rjg2_tetradodon n
124	16	50.0	206	2	04SKY8_TETNG	04sky8_tetradodon n	197	16	50.0	296	2	07X9B9_ARATH	07x9b9_arabidopsis
125	16	50.0	213	2	074ZG8_ASHGO	074zg8_ashbya goss	198	16	50.0	298	2	08H5F5_ORYSA	08h5f5_oryza sativ
126	16	50.0	214	2	09W3F3_DROME	09w3f3_drosophila	199	16	50.0	300	2	04SUT2_TETNG	04sut2_tetradodon n
127	16	50.0	215	2	07R4Z8_GIALA	07r4z8_giardia lam	200	16	50.0	306	2	084566_CHLTR	084566_chlamydia t
128	16	50.0	215	2	06QUH4_9GAMM	06quh4_methylomona	201	16	50.0	309	2	04STZ8_TETNG	04stz8_tetradodon n
129	16	50.0	215	2	06QUH6_9GAMM	06quh6_methylomona	202	16	50.0	311	2	051PB5_MAGGR	051pb5_magnaporthe
130	16	50.0	216	2	05TV10_ANOGA	05tv10_anopheles g	203	16	50.0	317	1	WRK15_ARATH	022176_arabidopsis
131	16	50.0	219	2	08G4A2_BIFLO	08g4a2_bifidobacte	204	16	50.0	321	2	09ZUS8_ARATH	09zus8_arabidopsis
132	16	50.0	218	2	063SB2_BURPS	063sb2_burkholderi	205	16	50.0	325	2	06AUJ4_ORYSA	06auj4_oryza sativ
133	16	50.0	221	2	05WNK7_CABER	05wnk7_caenorhabdi	206	16	50.0	335	2	0727E2_DESVH	0727e2_deuileovbr
134	16	50.0	222	2	04VSY6_DROME	04vys6_drosophila	207	16	50.0	340	2	07X1L8_ORYSA	07x1l8_oryza sativ
135	16	50.0	225	2	027014_TENMO	027014_tenebrio mo	208	16	50.0	346	2	07ZP51_9H1V1	07zps1_human immun
136	16	50.0	225	2	027013_TENMO	027013_tenebrio mo	209	16	50.0	346	2	04TWI8_9H1V1	04twi8_human immun
137	16	50.0	228	2	04V5S4_DROME	04v5s4_drosophila	210	16	50.0	347	2	04QCP8_LEIMA	04qcp8_leishmania
138	16	50.0	232	2	06ZLR3_BURMA	06zlr3_burkholderi	211	16	50.0	349	1	PAX4_RAT	088436_rattus norv
139	16	50.0	233	2	08LHK2_ORYSA	08lhk2_oryza sativ	212	16	50.0	349	2	04TWU5_9H1V1	04twu5_human immun
140	16	50.0	237	2	06YU04_ORYSA	06yu04_oryza sativ	213	16	50.0	351	2	04TWU9_9H1V1	04twu9_human immun
141	16	50.0	238	2	07QU60_GIALA	07qu60_giardia lam	214	16	50.0	353	2	06TVP8_ORYSA	06tvp8_oryza sativ
142	16	50.0	240	2	093877_FUSOX	093877_fusarium ox	215	16	50.0	353	2	04TWU7_9H1V1	04twu7_human immun
143	16	50.0	241	2	09VW4_DROME	09vww4_drosophila	216	16	50.0	354	2	04TWV0_9H1V1	04twv0_human immun
144	16	50.0	242	2	004701_NECHA	004701_nectria hae	217	16	50.0	355	2	08M1I0_BOVIN	08m1i0_bos tauru
145	16	50.0	243	2	05QHS3_BRARE	05qhs3_brachydanio	218	16	50.0	356	2	04TW2_9H1V1	04tw2_human immun
146	16	50.0	245	1	CTRB_GADMO	06046_gadus moriu	219	16	50.0	358	2	053KJ7_ORYSA	053kj7_oryza sativ
147	16	50.0	247	2	018458_HENGL	018458_heterodera	220	16	50.0	362	2	08CSN9_XANOR	08csn9_mus muscu
148	16	50.0	249	2	05VRX0_ORYSA	05vrx0_oryza sativ	221	16	50.0	363	2	08H1J8_HANOR	08h1j8_xanthomonas
149	16	50.0	251	2	024774_ENGBU	024774enchytraeus	222	16	50.0	366	2	05LUI3_SILPO	05lui3_elicibacte
150	16	50.0	251	2	0701Q4_ENGBU	0701q4_enchytraeus	223	16	50.0	370	2	07XEN8_ORYSA	07xen8_oryza sativ
151	16	50.0	252	1	TRYI_DROME	P52905_drosophila	224	16	50.0	371	2	04P4V1_USTMA	04p4v1_ustilego ma
152	16	50.0	254	2	0540Z6_DROME	0540z6_drosophila	225	16	50.0	373	2	06ABK7_PROAC	06abk7_proplionibac
153	16	50.0	254	2	060K89_CABER	060k89_caenorhabdi	226	16	50.0	373	2	05SNS5_BRARE	05sns5_brachydanio
154	16	50.0	258	2	09YOB9_DROME	09yob9_drosophila	227	16	50.0	374	2	07PBH7_ANOGA	07pbh7_anopheles g
155	16	50.0	259	2	09GQ40_GIALA	09gq40_giardia lam	228	16	50.0	374	2	07Z1R1_CABER	07z1r1_caenorhabdi
156	16	50.0	259	2	058PZ0_ARATH	058p20_arabidopsis	229	16	50.0	375	2	094699_PAPPR	094699_parametium
157	16	50.0	259	2	08H1N1_ARATH	08h1n1_arabidopsis	230	16	50.0	376	2	064CRL_9ARCH	064crl_uncultured
158	16	50.0	259	2	06AZC2_BRARE	06azc2_brachydanio	231	16	50.0	377	2	057WMD_9RRYP	057wd2_crypanosoma
159	16	50.0	260	2	06P2V9_XENTR	06p2v9_xenopus tro	232	16	50.0	378	2	063RGO_BURPS	063rgo_burkholderi
160	16	50.0	260	2	09W7Q3_PAROL	09w7q3_xenopus tro	233	16	50.0	378	2	063166_BURMA	063166_burkholderi
161	16	50.0	261	2	06DHD9_BRARE	06dhd9_brachydanio	234	16	50.0	380	2	09T0G1_ARATH	09t0g1_arabidopsis
162	16	50.0	261	2	07Z1I3_9H1V1	07z1i3_human immun	235	16	50.0	380	2	08XN74_CLOPE	08xn74_clostridium
163	16	50.0	262	1	VME1_CVCAI	P36299_canine ente	236	16	50.0	382	2	04W993_ASPFU	04w993_aspergillus
164	16	50.0	262	1	053KM3_ORYSA	053km3_oryza sativ	237	16	50.0	382	2	04QBZ6_LEIMA	04qbz6_leishmania
165	16	50.0	263	1	CTRA_GADMO	P47796_gadus moriu	238	16	50.0	386	2	06Z7H7_CABER	06z7h7_caenorhabdi
166	16	50.0	263	2	09PWQ6_GADMO	09pw6_gadus moriu	239	16	50.0	388	2	06P6D1_HUMAN	06p6d1_homo sapien
167	16	50.0	263	2	04SBW1_TETNG	04sbw1_tetradodon n	240	16	50.0	392	2	09P646_MOUSE	09p646_mus muscu
168	16	50.0	263	2	05H2D0_XENTR	05h2d0_xenopus tro	241	16	50.0	394	2	09GQ47_GIALA	09gq47_giardia lam
169	16	50.0	263	2	05I029_XENTR	05i029_xenopus tro	242	16	50.0	397	2	06Z3N1_CABER	06z3n1_caenorhabdi
170	16	50.0	263	2	06GNF7_XENTLA	06gnf7_xenopus lae	243	16	50.0	397	2	09XUB6_CABER	09xub6_caenorhabdi
171	16	50.0	263	2	06PGS4_XENTLA	06pgs4_xenopus lae	244	16	50.0	399	2	07SETE_NEUCR	07sete_neurospora
172	16	50.0	263	2	07SY84_XENTLA	07sy84_xenopus lae	245	16	50.0	400	2	063R93_BURPS	063r93_burkholderi
173	16	50.0	264	2	086BC4_SCHJA	086bc4_schistosoma	246	16	50.0	401	2	07PV94_ANOGA	07pv94_anopheles g
174	16	50.0	264	2	04QY78_SPAUV	04qy78_sparus aura	247	16	50.0	404	2	09C211_NEUCR	09c211_neurospora
175	16	50.0	264	2	0561U4_BRARE	0561u4_brachydanio	248	16	50.0	404	2	04RPNO_TETNG	04rpn0_tetradodon n
176	16	50.0	266	2	05W869_LEGPA	05w869_legionella	249	16	50.0	410	2	05HZU4_XENTR	05hzu4_xenopus tro
177	16	50.0	271	2	04S096_TETNG	04s096_tetradodon n	250	16	50.0	411	2	09SSV1_DROME	09ssv1_drosophila

251	16	50.0	411	2	Q9VW5	DROME	Q9VW5	drosophila	324	16	50.0	572	2	O42728	RHOCTO	O42728	rhodospirid
252	16	50.0	415	2	Q5KC03	CRYNB	Q5KC03	cryptococcu	325	16	50.0	574	1	GAGJ	DROFY	P21331	drosophila
253	16	50.0	415	2	Q5TP22	ANOGEA	Q5TP22	anopheles g	326	16	50.0	574	2	O5XTQ4	BOTCI	P52XQ4	botrytis ci
254	16	50.0	415	2	Q7V960	PROMM	Q7V960	prochloroco	327	16	50.0	574	2	O4R8L1	MAGFA	O4R8L1	macaca faac
255	16	50.0	417	2	Q4Q802	LEIMA	Q4Q802	leishmania	328	16	50.0	581	2	Q96WM9	BOTCI	Q96WM9	botrytis ci
256	16	50.0	420	2	Q4Q022	LEIMA	Q4Q022	leishmania	329	16	50.0	587	2	Q7RIZ0	GIALA	Q7RIZ0	giardia lam
257	16	50.0	421	1	Q6W340	9HIV1	Q6W340	human immun	330	16	50.0	589	1	Q6GQ06	MOUSE	Q6GQ06	mus musculu
258	16	50.0	422	1	RAD52	CHICK	RAD52	gallus gall	331	16	50.0	590	1	IRAK2	HUMAN	IRAK2	homo sapien
259	16	50.0	426	2	Q72M52	LEBIC	Q72M52	leptospiira	332	16	50.0	600	1	EGFL5	HUMAN	EGFL5	homo sapien
260	16	50.0	426	2	Q8F9T9	LEBIC	Q8F9T9	leptospiira	333	16	50.0	603	2	Q585K6	9TRYP	Q585K6	trypanosoma
261	16	50.0	430	2	Q840C2	9GMM	Q840C2	celivibrio	334	16	50.0	607	2	Q4JHG3	9RAMI	Q4JHG3	perilla cit
262	16	50.0	436	2	Q26859	TRYCO	Q26859	trypanosoma	335	16	50.0	608	2	Q5U388	BRARE	Q5U388	brachydanio
263	16	50.0	440	2	Q941N4	AVBSA	Q941N4	avena sativ	336	16	50.0	608	2	Q57UA7	9TRYP	Q57UA7	trypanosoma
264	16	50.0	441	2	Q6UAP5	TEFNG	Q6UAP5	tetradodon n	337	16	50.0	613	2	Q95TQ2	DROME	Q95TQ2	drosophila
265	16	50.0	443	2	Q6AZB6	BRARE	Q6AZB6	brachydanio	338	16	50.0	613	2	Q9V5F6	DROME	Q9V5F6	drosophila
266	16	50.0	446	2	Q7XGP0	ORYSA	Q7XGP0	oryza sativ	339	16	50.0	616	2	Q738S5	BACCI	Q738S5	bacillus ce
267	16	50.0	446	2	Q8SB64	ORYSA	Q8SB64	oryza sativ	340	16	50.0	622	2	Q7QWD2	GIALA	Q7QWD2	giardia lam
268	16	50.0	447	2	Q8TL45	METAC	Q8TL45	methanobarc	341	16	50.0	625	2	Q5K546	HUMAN	Q5K546	homo sapien
269	16	50.0	448	2	Q5GVX5	XANOR	Q5GVX5	xanthomonas	342	16	50.0	625	2	Q5R810	POMVM	Q5R810	pongo pygma
270	16	50.0	450	2	Q7XP30	ORYSA	Q7XP30	oryza sativ	343	16	50.0	629	2	Q9DWM3	RCWVM	Q9DWM3	rat cytohes
271	16	50.0	452	2	Q8Q000	METMA	Q8Q000	methanobarc	344	16	50.0	632	2	Q6YXW2	ORYSA	Q6YXW2	oryza sativ
272	16	50.0	453	2	Q20490	CABEL	Q20490	caenorhabdi	345	16	50.0	633	2	Q8WTK1	9HIV1	Q8WTK1	meloidogyne
273	16	50.0	455	2	Q4PT10	ARATH	Q4PT10	arabidopsis	346	16	50.0	636	2	Q61BP2	CABER	Q61BP2	caenorhabdi
274	16	50.0	455	2	Q61FV6	RAT	Q61FV6	rattus norv	347	16	50.0	638	1	Y132	CABEL	Y132	CABEL
275	16	50.0	457	2	Q9YQ21	9HERP	Q9YQ21	randid herpe	348	16	50.0	639	2	Q84K89	AVESA	Q84K89	avena sativ
276	16	50.0	461	2	Q4Q6C9	LEIMA	Q4Q6C9	leishmania	349	16	50.0	642	2	Q9PPH1	CAMJE	Q9PPH1	campylobact
277	16	50.0	463	2	Q6Z3D6	ORYSA	Q6Z3D6	oryza sativ	350	16	50.0	645	2	Q97448	GIALA	Q97448	giardia lam
278	16	50.0	464	2	Q5CIB1	CRYPV	Q5CIB1	cryptospori	351	16	50.0	652	2	Q4S9Q3	TEFNG	Q4S9Q3	tetradodon n
279	16	50.0	464	2	Q5CIV7	CRYHO	Q5CIV7	cryptospori	352	16	50.0	660	2	Q7QY47	GIALA	Q7QY47	giardia lam
280	16	50.0	465	2	Q7QVC3	GIALA	Q7QVC3	giardia lam	353	16	50.0	660	2	Q6FCHO	XENLA	Q6FCHO	xenopus lae
281	16	50.0	466	2	Q5U8X1	9HIV1	Q5U8X1	human immun	354	16	50.0	665	1	YDER	SCHPO	YDER	SCHPO
282	16	50.0	468	2	Q8GY60	ARATH	Q8GY60	arabidopsis	355	16	50.0	667	2	Q9XTK3	GIALA	Q9XTK3	giardia lam
283	16	50.0	470	2	Q82BGO	STRMA	Q82BGO	streptomyce	356	16	50.0	681	2	Q6PUC7	ANOGA	Q6PUC7	anopheles g
284	16	50.0	471	2	Q5YTV3	NOCFA	Q5YTV3	nocardia fa	357	16	50.0	681	2	Q6LBN7	BOVIN	Q6LBN7	bos taurus
285	16	50.0	473	2	Q5ZCU2	MAGGR	Q5ZCU2	magnaporthe	358	16	50.0	687	2	Q817G1	CABEL	Q817G1	caenorhabdi
286	16	50.0	473	2	Q5PAD8	NEIG1	Q5PAD8	neisseria g	359	16	50.0	695	1	TRFL	HORSE	TRFL	HORSE
287	16	50.0	473	2	Q9JVM9	NEIMA	Q9JVM9	neisseria m	360	16	50.0	696	2	Q5R2V7	HUMAN	Q5R2V7	homo sapien
288	16	50.0	473	2	Q9JXZ7	NEIMB	Q9JXZ7	neisseria m	361	16	50.0	699	2	Q93JN9	STRAM	Q93JN9	streptomyce
289	16	50.0	475	2	Q9X8J5	STRCO	Q9X8J5	streptomyce	362	16	50.0	708	1	TRFL	BOVIN	TRFL	BOVIN
290	16	50.0	479	2	Q4SLI6	TEFNG	Q4SLI6	tetradodon n	363	16	50.0	708	1	TRFL	BUBBU	TRFL	BUBBU
291	16	50.0	480	2	Q7SZR0	BRARE	Q7SZR0	brachydanio	364	16	50.0	708	1	TRFL	CAMDR	TRFL	CAMDR
292	16	50.0	481	2	Q9D635	MOUSE	Q9D635	mus musculu	365	16	50.0	708	1	Q61EC7	BOVIN	Q61EC7	bos taurus
293	16	50.0	486	2	Q91Q21	ARATH	Q91Q21	arabidopsis	366	16	50.0	711	1	TRFL	HUMAN	TRFL	HUMAN
294	16	50.0	488	2	Q5WMJ6	LEGPL	Q5WMJ6	legionella	367	16	50.0	711	1	Q4QD78	LEIMA	Q4QD78	leishmania
295	16	50.0	489	2	Q5AL61	CANAL	Q5AL61	candida alb	368	16	50.0	711	2	Q4QD79	LEIMA	Q4QD79	leishmania
296	16	50.0	496	2	Q4MYK7	9DBTL	Q4MYK7	anaeromyxob	369	16	50.0	711	2	Q5DSMO	HUMAN	Q5DSMO	homo sapien
297	16	50.0	501	2	Q4HVA7	GIIZE	Q4HVA7	glibereella	370	16	50.0	713	2	Q6FES7	ACTAD	Q6FES7	actinobact
298	16	50.0	510	1	MS2P	CRIGR	MS2P	azotobacter	371	16	50.0	719	1	Q75715	ASHGO	Q75715	ashbya goss
299	16	50.0	513	2	Q41VN9	AZQVI	Q41VN9	azotobacter	372	16	50.0	722	2	Q6FR22	CANCA	Q6FR22	candida gla
300	16	50.0	517	2	Q4VBH3	RAT	Q4VBH3	rattus norv	373	16	50.0	734	2	Q9NSE6	HUMAN	Q9NSE6	homo sapien
301	16	50.0	520	2	Q8NVC0	HUMAN	Q8NVC0	homo sapien	374	16	50.0	740	2	Q9UPP2	BOVIN	Q9UPP2	homo sapien
302	16	50.0	520	2	Q81YS4	HUMAN	Q81YS4	homo sapien	375	16	50.0	741	2	Q57VP8	9TRYP	Q57VP8	trypanosoma
303	16	50.0	523	2	Q53P11	ORYSA	Q53P11	oryza sativ	376	16	50.0	743	2	Q4QD80	LEIMA	Q4QD80	leishmania
304	16	50.0	528	2	Q86MK1	DROME	Q86MK1	drosophila	377	16	50.0	777	2	Q4QD78	LEIMA	Q4QD78	leishmania
305	16	50.0	532	2	Q53NH3	ORYSA	Q53NH3	oryza sativ	378	16	50.0	777	2	Q4QD79	LEIMA	Q4QD79	leishmania
306	16	50.0	532	2	Q8X1C0	USTMA	Q8X1C0	ustilago ma	379	16	50.0	783	2	Q9PBVA	STRCO	Q9PBVA	streptomyce
307	16	50.0	533	2	Q4P6V3	USTMA	Q4P6V3	ustilago ma	380	16	50.0	783	2	Q76128	9VIRU	Q76128	heilicobasid
308	16	50.0	533	2	Q7ABC9	GBOSL	Q7ABC9	geobacter b	381	16	50.0	783	2	Q5NMW2	9ARCH	Q5NMW2	uncultured
309	16	50.0	536	2	Q4U3U6	NEUAF	Q4U3U6	neutrospora	382	16	50.0	795	2	Q4TEN4	GIIZE	Q4TEN4	glibereella
310	16	50.0	536	2	Q4U3U4	9PEZI	Q4U3U4	neutrospora	383	16	50.0	802	2	Q4ID36	GIIZE	Q4ID36	glibereella
311	16	50.0	537	1	Q4YBUX5	TEFNG	Q4YBUX5	tetradodon n	384	16	50.0	810	2	Q5VO42	ORYSA	Q5VO42	oryza sativ
312	16	50.0	537	1	QAY	NEUCR	QAY	neurospora	385	16	50.0	825	2	Q4Q608	LEIMA	Q4Q608	leishmania
313	16	50.0	538	2	Q6Z1B7	ORYSA	Q6Z1B7	aspergillus	386	16	50.0	825	2	Q7RFP7	PLAYO	Q7RFP7	plasmodium
314	16	50.0	542	2	Q9FFP95	ARATH	Q9FFP95	arabidopsis	387	16	50.0	842	2	Q61UW9	CABER	Q61UW9	caenorhabdi
315	16	50.0	543	2	Q86ZM8	PODAN	Q86ZM8	podospora a	388	16	50.0	843	2	Q8B124	XANAC	Q8B124	xanthomonas
316	16	50.0	545	2	Q4IAQ1	GIIZE	Q4IAQ1	glibereella	389	16	50.0	845	2	Q8AK00	9HIV1	Q8AK00	human immun
317	16	50.0	546	2	Q5M281	9AGAR	Q5M281	pleurotus s	390	16	50.0	845	2	Q4Q607	LEIMA	Q4Q607	leishmania
318	16	50.0	549	2	Q51XV6	MAGGR	Q51XV6	magnaporthe	391	16	50.0	847	2	Q4Q609	LEIMA	Q4Q609	leishmania
319	16	50.0	552	2	Q4IRQ1	GIIZE	Q4IRQ1	glibereella	392	16	50.0	847	2	Q8AF19	9HIV1	Q8AF19	human immun
320	16	50.0	552	2	Q6DDA9	XENTR	Q6DDA9	xenopus tro	393	16	50.0	848	2	Q4QWTS	9HIV1	Q4QWTS	human immun
321	16	50.0	560	2	Q5B514	EMENI	Q5B514	aspergillus	394	16	50.0	851	2	Q7RVV6	NEUCR	Q7RVV6	neutrospora
322	16	50.0	561	2	Q96WM0	BOTCI	Q96WM0	botrytis ci	395	16	50.0	852	2				
323	16	50.0	561	2	Q96WM0	BOTCI	Q96WM0	botrytis ci	396	16	50.0	852	2				

397	16	50.0	852	2	Q9HE44_NEUCR	Q9HE44_neutropora	470	16	50.0	2128	2	Q5ONG8_EWYHI	Q5ONG8_entamoeba h
398	16	50.0	852	2	Q70010_9H1V1	Q70010 human immun	471	16	50.0	2233	2	Q94711_PAK1B	Q94711_parametium
399	16	50.0	852	2	Q5BEC9_9H1V1	Q5BEC9 human immun	472	16	50.0	2273	2	Q5ICB9_MACEA	Q5ICB9_maceca fasc
400	16	50.0	858	2	Q8AE09_9H1V1	Q8AE09 human immun	473	16	50.0	2377	2	Q7W080_BORBR	Q7W080_boderella
401	16	50.0	859	2	Q4QG27_LEITMA	Q4QG27 leishmania	474	16	50.0	2397	2	Q27167_PARTE	Q27167_parametium
402	16	50.0	859	2	Q8ADK7_9H1V1	Q8ADK7 human immun	475	16	50.0	2395	2	Q671Y4_PAPPR	Q671Y4_parametium
403	16	50.0	859	2	Q8UMP8_9H1V1	Q8UMP8 human immun	476	16	50.0	2533	2	P90569_PARTE	P90569_parametium
404	16	50.0	860	2	Q998E5_9H1V1	Q998E5 human immun	477	16	50.0	2533	2	Q27183_PARTE	Q27183_parametium
405	16	50.0	864	2	Q4U536_9H1V1	Q4U536 human immun	478	16	50.0	2543	2	P90649_PAPPR	P90649_parametium
406	16	50.0	865	2	Q73342_9H1V1	Q73342 human immun	479	16	50.0	2715	1	G156_PAPPR	G156_parametium
407	16	50.0	865	2	Q73353_9H1V1	Q73353 human immun	480	16	50.0	2986	1	Q4HMU4_GIBZE	Q4HMU4_gibberella
408	16	50.0	870	2	Q5BED7_9H1V1	Q5BED7 human immun	481	16	50.0	3596	1	REOA_SHFV	REOA772 s replicase
409	16	50.0	877	2	Q5YGT1_9H1V1	Q5YGT1 human immun	482	16	50.0	3869	2	Q5RCRO_CRYPV	Q5RCRO_cryptospori
410	16	50.0	879	2	Q4W9A0_ASPFU	Q4W9A0 aspergillus	483	16	50.0	3869	2	Q8BPQ3_CRYPV	Q8BPQ3_cryptospori
411	16	50.0	879	2	Q4PGC4_USTMA	Q4PGC4 ustilago ma	484	16	50.0	4382	2	Q5DIC6_CVHSA	Q5DIC6_sars corona
412	16	50.0	881	2	Q4N476_THBPA	Q4N476 theileria p	485	16	50.0	4382	2	Q5Y188_CVHSA	Q5Y188_sars corona
413	16	50.0	884	2	Q80718_9H1V1	Q80718 human immun	486	16	50.0	4382	2	Q692E5_CVHSA	Q692E5_sars corona
414	16	50.0	889	2	Q81Y03_PLAF7	Q81Y03 plasmodium	487	16	50.0	4382	2	Q6JH39_CVHSA	Q6JH39_sars corona
415	16	50.0	897	2	Q5SS04_CRYNE	Q5SS04 cryptococcu	488	16	50.0	4382	2	Q6JH47_CVHSA	Q6JH47_sars corona
416	16	50.0	897	2	Q5KGF6_CRYNE	Q5KGF6 cryptococcu	489	16	50.0	4382	2	Q6R717_CVHSA	Q6R717_sars corona
417	16	50.0	907	2	Q9FW49_ARATH	Q9FW49 arabidopsis	490	16	50.0	4382	2	Q6RCW6_CVHSA	Q6RCW6_sars corona
418	16	50.0	911	2	Q80TJ8_MOUSE	Q80TJ8 mus musculi	491	16	50.0	4382	2	Q6RCX7_CVHSA	Q6RCX7_sars corona
419	16	50.0	912	2	Q4JHL4_FUGRU	Q4JHL4 fuga rubrip	492	16	50.0	4382	2	Q6RCY8_CVHSA	Q6RCY8_sars corona
420	16	50.0	918	2	Q6E7C9_GIBOC	Q6E7C9 cikopleure	493	16	50.0	4382	2	Q6RCZ9_CVHSA	Q6RCZ9_sars corona
421	16	50.0	920	2	Q7XGN4_ORISA	Q7XGN4 oryza sativ	494	16	50.0	4382	2	Q6RD10_CVHSA	Q6RD10_sars corona
422	16	50.0	920	2	Q7XGN4_ORISA	Q7XGN4 oryza sativ	495	16	50.0	4382	2	Q6RD21_CVHSA	Q6RD21_sars corona
423	16	50.0	929	2	Q5ERD8_9H1V1	Q5ERD8 human immun	496	16	50.0	4382	2	Q6RD32_CVHSA	Q6RD32_sars corona
424	16	50.0	934	2	Q87R90_VIBPA	Q87R90 vibrio para	497	16	50.0	4382	2	Q6RD43_CVHSA	Q6RD43_sars corona
425	16	50.0	935	2	Q4P9H9_USTMA	Q4P9H9 ustilago ma	498	16	50.0	4382	2	Q6RD54_CVHSA	Q6RD54_sars corona
426	16	50.0	939	2	Q5RRH6_BRARE	Q5RRH6 brachydanio	499	16	50.0	4382	2	Q6RD65_CVHSA	Q6RD65_sars corona
427	16	50.0	967	2	Q4T490_TERNG	Q4T490 tetradodon n	500	16	50.0	4382	2	Q6VA79_CVHSA	Q6VA79_sars corona
428	16	50.0	975	2	P91357_CABEL	P91357 caenorhabdi	501	16	50.0	4382	2	Q6VA90_CVHSA	Q6VA90_sars corona
429	16	50.0	999	1	SMG_DROME	Q23972 dirosophila	502	16	50.0	4382	2	Q6VAA1_CVHSA	Q6VAA1_sars corona
430	16	50.0	1031	2	Q65500_ARATH	Q65500 arabidopsis	503	16	50.0	4382	2	Q6WGN0_CVHSA	Q6WGN0_sars corona
431	16	50.0	1038	2	Q5B3N6_9TRYP	Q5B3N6 trypanosoma	504	16	50.0	5560	1	SPEN_DROME	SPEN_DROME
432	16	50.0	1039	2	Q4P295_USTMA	Q4P295 ustilago ma	505	16	50.0	6880	1	Q6S8D8_CVHSA	Q6S8D8_sars corona
433	16	50.0	1069	2	Q532X1_CANAL	Q532X1 candida alb	506	16	50.0	7073	1	RIAB_CVHSA	RIAB_CVHSA
434	16	50.0	1069	2	Q9J3F2_CABEL	Q9J3F2 caenorhabdi	507	16	50.0	7073	2	Q6DIC7_CVHSA	Q6DIC7_sars corona
435	16	50.0	1071	2	Q61ZG1_CABEL	Q61ZG1 caenorhabdi	508	16	50.0	7073	2	Q5Y189_CVHSA	Q5Y189_sars corona
436	16	50.0	1105	2	Q20371_CABEL	Q20371 caenorhabdi	509	16	50.0	7073	2	Q692E6_CVHSA	Q692E6_sars corona
437	16	50.0	1136	2	Q92C14_LISIN	Q92C14 listeria in	510	16	50.0	7073	2	Q6JH40_CVHSA	Q6JH40_sars corona
438	16	50.0	1154	2	Q9G046_GIALA	Q9G046 giardia lam	511	16	50.0	7073	2	Q6JH48_CVHSA	Q6JH48_sars corona
439	16	50.0	1168	2	Q6AVL1_ORISA	Q6AVL1 oryza sativ	512	16	50.0	7073	2	Q6R7Y8_CVHSA	Q6R7Y8_sars corona
440	16	50.0	1191	2	Q8TOL3_DROME	Q8TOL3 dirosophila	513	16	50.0	7073	2	Q6TZF1_CVHSA	Q6TZF1_sars corona
441	16	50.0	1194	2	Q4FX33_LEITMA	Q4FX33 leishmania	514	16	50.0	7073	2	Q6UZF1_CVHSA	Q6UZF1_sars corona
442	16	50.0	1194	2	Q76M68_RAT	Q76M68 rattus norv	515	16	50.0	7073	2	Q6UZE5_CVHSA	Q6UZE5_sars corona
443	16	50.0	1227	2	Q6IFU2_SCHMA	Q6IFU2 schistosoma	516	16	50.0	7073	2	Q6V5B6_CVHSA	Q6V5B6_sars corona
444	16	50.0	1243	2	Q8OPT3_WSSV	Q8OPT3 white spot	517	16	50.0	7073	2	Q6RCW7_CVHSA	Q6RCW7_sars corona
445	16	50.0	1243	2	Q8VAP5_WSSV	Q8VAP5 white spot	518	16	50.0	7073	2	Q6RCX8_CVHSA	Q6RCX8_sars corona
446	16	50.0	1243	2	Q91LM1_WSSV	Q91LM1 white spot	519	16	50.0	7073	2	Q6RCY9_CVHSA	Q6RCY9_sars corona
447	16	50.0	1270	2	Q5UXR6_HUMAN	Q5UXR6 homo sapien	520	16	50.0	7073	2	Q6RD00_CVHSA	Q6RD00_sars corona
448	16	50.0	1310	1	GPI25_MOUSE	Q7C36 mus musculi	521	16	50.0	7073	2	Q6RD11_CVHSA	Q6RD11_sars corona
449	16	50.0	1318	2	Q4ODAS_LEITMA	Q4ODAS leishmania	522	16	50.0	7073	2	Q6RD22_CVHSA	Q6RD22_sars corona
450	16	50.0	1332	2	Q4FPY5_LEITMA	Q4FPY5 leishmania	523	16	50.0	7073	2	Q6RD33_CVHSA	Q6RD33_sars corona
451	16	50.0	1353	2	Q4QBR1_LEITMA	Q4QBR1 leishmania	524	16	50.0	7073	2	Q6RD44_CVHSA	Q6RD44_sars corona
452	16	50.0	1363	2	Q4L1W7_9BPRK	Q4L1W7 burkholderi	525	16	50.0	7073	2	Q6RD55_CVHSA	Q6RD55_sars corona
453	16	50.0	1367	2	Q4QAT6_LEITMA	Q4QAT6 leishmania	526	16	50.0	7073	2	Q6RD66_CVHSA	Q6RD66_sars corona
454	16	50.0	1369	2	Q4QAT7_LEITMA	Q4QAT7 leishmania	527	16	50.0	7073	2	Q6VA80_CVHSA	Q6VA80_sars corona
455	16	50.0	1380	1	CYAA_LEITMA	Q27675 leishmania	528	16	50.0	7073	2	Q6VA91_CVHSA	Q6VA91_sars corona
456	16	50.0	1467	2	Q61ON6_CABEL	Q61ON6 caenorhabdi	529	16	50.0	7073	2	Q6VAA2_CVHSA	Q6VAA2_sars corona
457	16	50.0	1474	2	Q62504_CABEL	Q62504 caenorhabdi	530	16	50.0	7073	2	Q4JDOS_CVHSA	Q4JDOS_sars corona
458	16	50.0	1486	2	Q4SDM0_TERNG	Q4SDM0 tetradodon n	531	16	50.0	7073	2	Q4JDR8_CVHSA	Q4JDR8_sars corona
459	16	50.0	1557	2	Q4QHE6_LEITMA	Q4QHE6 leishmania	532	16	46.9	26	2	Q93940_PODAN	Q93940 podospora a
460	16	50.0	1558	2	Q55TP3_CRYNE	Q55TP3 cryptococcu	533	15	46.9	33	2	Q9BYS0_HUMAN	Q9BYS0 homo sapien
461	16	50.0	1588	2	Q5KIP2_CRYNE	Q5KIP2 cryptococcu	534	15	46.9	40	1	MT1_DROME	MT1_DROME
462	16	50.0	1755	2	Q4QOPD_LEITMA	Q4QOPD leishmania	535	15	46.9	40	1	MT1_DROME	MT1_DROME
463	16	50.0	1769	2	Q7WC76_BORBP	Q7WC76 boderella	536	15	46.9	40	1	MT1_DROS1	MT1_DROS1
464	16	50.0	1818	1	BNC_DROME	Q8mxi dirosophila	537	15	46.9	43	1	COAT2_BPRHL	COAT2_bacterioph
465	16	50.0	1823	2	Q5B1H9_DROME	Q5b1h9 dirosophila	538	15	46.9	43	1	COAT2_XANCP	COAT2_xanthomonas
466	16	50.0	1823	2	Q7KV61_DROME	Q7KV61 dirosophila	539	15	46.9	43	2	Q4UUV5_XANCP	Q4UUV5_xanthomonas
467	16	50.0	2071	2	Q7QV58_GIALA	Q7QV58 giardia lam	540	15	46.9	43	2	Q4UUV5_XANCP	Q4UUV5_xanthomonas
468	16	50.0	2087	2	Q4S488_TERNG	Q4S488 tetradodon n	541	15	46.9	46	2	Q4UUV5_XANCP	Q4UUV5_xanthomonas
469	16	50.0	2117	2	Q4QIS6_LEITMA	Q4QIS6 leishmania	542	15	46.9	46	2	Q8PDK1_XANCP	Q8PDK1_xanthomonas



543	15	46.9	53	2	Q69338	9GAMA	Q69338	saimiri	line	616	15	46.9	101	2	009129	MOUSE	009129	mus	musculus
544	15	46.9	54	2	Q749S9	GEOSL	Q749S9	geobacter	s	617	15	46.9	101	2	Q71260	9H1V1	Q71260	human	immun
545	15	46.9	56	2	Q5C2S7	SCHJA	Q5C2S7	schistosoma		618	15	46.9	101	2	Q90DH4	9H1V1	Q90DH4	human	immun
546	15	46.9	57	2	Q60CB8	MERCA	Q60CB8	methylococc		619	15	46.9	101	2	Q90DH7	9H1V1	Q90DH7	human	immun
547	15	46.9	58	2	Q96914	9BETA	Q96914	human	hepce	620	15	46.9	102	1	VP07	BPAPS	VP07	bacterioph	
548	15	46.9	58	2	Q96915	9BETA	Q96915	human	hepce	621	15	46.9	102	2	Q6MNI21	BDBBA	Q6MNI21	human	immun
549	15	46.9	58	2	Q4TVJ8	TEING	Q4TVJ8	tetradon	n	622	15	46.9	102	2	Q71271	9H1V1	Q71271	human	immun
550	15	46.9	59	2	Q7S2A3	NEUCR	Q7S2A3	neuropora		623	15	46.9	102	2	Q90DI4	9H1V1	Q90DI4	human	immun
551	15	46.9	59	2	Q5CINA	SCHJA	Q5CINA	schistosoma		624	15	46.9	103	2	Q7ORX6	GIATA	Q7ORX6	giardia	lam
552	15	46.9	59	2	Q4TRJ3	TEING	Q4TRJ3	tetradon	n	625	15	46.9	103	2	Q9UOT5	LOCOMI	Q9UOT5	loocusta	mtg
553	15	46.9	61	2	Q86TX3	HUMAN	Q86TX3	homo	sepien	626	15	46.9	103	2	Q86RT2	9HEPC	Q86RT2	hepatitis	c
554	15	46.9	63	2	Q6FJN6	CANGA	Q6FJN6	canidia	glia	627	15	46.9	104	2	Q7PG48	ANOCA	Q7PG48	anopheles	g
555	15	46.9	64	2	Q6KOM3	9VIRU	Q6KOM3	human	astro	628	15	46.9	104	2	Q6W9S3	ORYSA	Q6W9S3	oryza	sativ
556	15	46.9	67	2	Q8JMX7	9VIRU	Q8JMX7	human	astro	629	15	46.9	104	2	Q84BA7	ERWCH	Q84BA7	erythrina	chr
557	15	46.9	67	2	Q4TGE5	TEING	Q4TGE5	tetradon	n	630	15	46.9	106	2	Q8NOX0	CORGL	Q8NOX0	corynebacte	
558	15	46.9	68	2	Q4THE0	TEING	Q4THE0	tetradon	n	631	15	46.9	107	2	Q7R3Y4	GIATA	Q7R3Y4	giardia	lam
559	15	46.9	72	2	Q5DBB6	ANOCA	Q5DBB6	anopheles	g	632	15	46.9	107	2	Q84NA3	9RABA	Q84NA3	zapoteca	te
560	15	46.9	73	2	Q5RIJ9	PIC	Q5RIJ9	sus	scrofa	633	15	46.9	107	2	Q9DZK3	9H1V1	Q9DZK3	human	immun
561	15	46.9	73	2	Q62127	BURMA	Q62127	burholderi		634	15	46.9	107	2	Q9DZK4	9H1V1	Q9DZK4	human	immun
562	15	46.9	74	2	Q8YSS5	ANASP	Q8YSS5	anabaena	sp	635	15	46.9	107	2	Q9DZK6	9H1V1	Q9DZK6	human	immun
563	15	46.9	74	2	Q5YF21	9VIRU	Q5YF21	rock	bream	636	15	46.9	107	2	Q9DZK5	9H1V1	Q9DZK5	human	immun
564	15	46.9	77	2	Q4UZT9	XANCP	Q4UZT9	xanthomonas		637	15	46.9	109	1	YGIC	YEAST	YGIC	yeast	
565	15	46.9	77	2	Q4UZT9	XANCP	Q4UZT9	xanthomonas		638	15	46.9	109	2	Q8CEA7	MOUSE	Q8CEA7	mus	musculus
566	15	46.9	78	2	Q8PDJ8	XANCP	Q8PDJ8	xanthomonas		639	15	46.9	109	2	Q8JG58	AMMBE	Q8JG58	ambystoma	m
567	15	46.9	78	2	Q06467	XANCA	Q06467	xanthomonas		640	15	46.9	109	2	Q6CTJ3	KLUTA	Q6CTJ3	kluyveromyc	
568	15	46.9	79	2	Q26086	9TURB	Q26086	polycellus	n	641	15	46.9	111	2	Q9ZHB2	CHLUP	Q9ZHB2	chlamydia	p
569	15	46.9	80	2	Q92MT4	RHIME	Q92MT4	rhizobium	m	642	15	46.9	111	2	Q8C5D5	MOUSE	Q8C5D5	mus	musculus
570	15	46.9	80	2	Q4J5G5	AZOV1	Q4J5G5	azotobacter		643	15	46.9	111	2	Q8C5D5	MOUSE	Q8C5D5	mus	musculus
571	15	46.9	80	2	Q5NCE5	MOUSE	Q5NCE5	mus	musculus	644	15	46.9	112	2	Q56714	9HEPC	Q56714	hepatitis	c
572	15	46.9	80	2	Q69752	9H1V1	Q69752	human	immun	645	15	46.9	112	2	Q56716	9HEPC	Q56716	hepatitis	c
573	15	46.9	81	2	Q80226	9H1V1	Q80226	human	immun	646	15	46.9	112	2	Q56717	9HEPC	Q56717	hepatitis	c
574	15	46.9	82	1	LCR45	ARATH	LCR45	arabidopsis		647	15	46.9	112	2	Q56718	9HEPC	Q56718	hepatitis	c
575	15	46.9	85	2	Q9HYT5	HUMAN	Q9HYT5	homo	sapien	648	15	46.9	112	2	Q56724	9HEPC	Q56724	hepatitis	c
576	15	46.9	85	2	Q84GQ2	9SYNE	Q84GQ2	synecchococ		649	15	46.9	112	2	Q56726	9HEPC	Q56726	hepatitis	c
577	15	46.9	85	2	P88625	9H1V1	P88625	human	immun	650	15	46.9	112	2	Q56727	9HEPC	Q56727	hepatitis	c
578	15	46.9	86	2	Q6TRY3	CULOU	Q6TRY3	culex	guinq	651	15	46.9	112	2	Q56728	9HEPC	Q56728	hepatitis	c
579	15	46.9	88	2	Q94203	CLAFU	Q94203	cladosporiu		652	15	46.9	112	2	Q56729	9HEPC	Q56729	hepatitis	c
580	15	46.9	88	2	Q90DH6	9H1V1	Q90DH6	human	immun	653	15	46.9	112	2	Q56730	9HEPC	Q56730	hepatitis	c
581	15	46.9	89	2	Q43185	HUMAN	Q43185	homo	sapien	654	15	46.9	112	2	Q56731	9HEPC	Q56731	hepatitis	c
582	15	46.9	89	2	Q51BM5	ENTHI	Q51BM5	entamoeba	h	655	15	46.9	112	2	Q90DH8	9H1V1	Q90DH8	human	immun
583	15	46.9	89	2	Q41UR5	AZOV1	Q41UR5	azotobacter		656	15	46.9	113	1	Y011	BPTA	Y011	bpt	
584	15	46.9	90	2	Q6TVM4	9POXV	Q6TVM4	orf	virus	657	15	46.9	113	1	Q5AJZ9	CANAL	Q5AJZ9	canidia	alb
585	15	46.9	91	2	Q4XC49	PLACH	Q4XC49	plasmodium		658	15	46.9	113	2	Q614J6	CABRB	Q614J6	caenorhabdi	
586	15	46.9	91	2	Q9SYI6	ARATH	Q9SYI6	arabidopsis		659	15	46.9	114	2	Q95ZX8	CABEL	Q95ZX8	caenorhabdi	
587	15	46.9	91	2	Q39862	9VIRU	Q39862	human	astro	660	15	46.9	114	2	Q90DI0	9H1V1	Q90DI0	human	immun
588	15	46.9	91	2	Q6TWM4	9POXV	Q6TWM4	orf	virus	661	15	46.9	115	2	Q9S9A1	VISAL	Q9S9A1	vicium	albu
589	15	46.9	92	2	Q9UIG9	HUMAN	Q9UIG9	homo	sapien	662	15	46.9	115	2	Q6RXU0	9H1V1	Q6RXU0	human	immun
590	15	46.9	92	2	Q8SAX8	ORYSA	Q8SAX8	oryza	sativ	663	15	46.9	116	2	Q4NZF5	9DELT	Q4NZF5	anaeromyxob	
591	15	46.9	92	2	P88627	9H1V1	P88627	human	immun	664	15	46.9	117	2	Q8UJ06	9H1V1	Q8UJ06	human	immun
592	15	46.9	93	2	Q5FMF0	IACAC	Q5FMF0	lacobacilli		665	15	46.9	117	2	Q8UJ07	9H1V1	Q8UJ07	human	immun
593	15	46.9	93	2	Q41098	CHYPL	Q41098	paramacium		666	15	46.9	117	2	Q8UJ08	9H1V1	Q8UJ08	human	immun
594	15	46.9	93	2	Q5FZJ8	9HEPC	Q5FZJ8	hepatitis	c	667	15	46.9	117	2	CG023	HUMAN	CG023	homo	sepien
595	15	46.9	93	2	Q5FZK1	9HEPC	Q5FZK1	hepatitis	c	668	15	46.9	118	1	Q8RFJ3	FUSNN	Q8RFJ3	fusobacteri	
596	15	46.9	93	2	Q5FZK8	9HEPC	Q5FZK8	hepatitis	c	669	15	46.9	118	2	Q7U993	SYNBP	Q7U993	synecchococ	
597	15	46.9	93	2	Q78021	9H1V1	Q78021	human	immun	670	15	46.9	119	2	Q8UJ09	9H1V1	Q8UJ09	human	immun
598	15	46.9	94	2	Q90DH5	9H1V1	Q90DH5	human	immun	671	15	46.9	120	2	Q69960	9H1V1	Q69960	human	immun
599	15	46.9	94	2	Q9PXN5	9H1V1	Q9PXN5	human	immun	672	15	46.9	121	2	Q69964	9H1V1	Q69964	human	immun
600	15	46.9	95	2	Q8CES2	MOUSE	Q8CES2	mus	musculus	673	15	46.9	121	2	YK4	CVHSA	YK4	cvhsa	
601	15	46.9	96	2	Q9BYU5	HUMAN	Q9BYU5	homo	sapien	674	15	46.9	122	1	Q5KQI2	ORYSA	Q5KQI2	oryza	sativ
602	15	46.9	96	2	Q7YTR6	CABEL	Q7YTR6	caenorhabdi		675	15	46.9	122	2	Q5X205	LEGPA	Q5X205	legionella	
603	15	46.9	96	2	Q8VD96	MESAU	Q8VD96	mesocricetu		676	15	46.9	122	2	Q5ZSJ7	LECPH	Q5ZSJ7	leishmania	
604	15	46.9	96	2	Q77367	9H1V1	Q77367	human	immun	677	15	46.9	122	2	Q82209	CHLUP	Q82209	chlamydomophi	
605	15	46.9	96	2	Q90DH9	9H1V1	Q90DH9	human	immun	678	15	46.9	122	2	Q5Y181	CVHSA	Q5Y181	sars	corona
606	15	46.9	97	2	Q8TFPH4	METAC	Q8TFPH4	methanosaarc		679	15	46.9	122	2	Q692D8	CVHSA	Q692D8	sars	corona
607	15	46.9	97	2	Q4Y107	PLABE	Q4Y107	plasmodium		680	15	46.9	122	2	Q6GYO3	CVHSA	Q6GYO3	sars	corona
608	15	46.9	98	2	Q4UZT6	XANCP	Q4UZT6	xanthomonas		681	15	46.9	122	2	Q6R7Y0	CVHSA	Q6R7Y0	sars	corona
609	15	46.9	98	2	Q8PDI3	XANCP	Q8PDI3	xanthomonas		682	15	46.9	122	2	Q6S8D1	CVHSA	Q6S8D1	sars	corona
610	15	46.9	98	2	Q90DI3	9H1V1	Q90DI3	human	immun	683	15	46.9	122	2	Q6S288	CVHSA	Q6S288	sars	corona
611	15	46.9	99	2	Q5Y4U4	AGBOR	Q5Y4U4	agelena	ori	684	15	46.9	122	2	Q6R1D6	CVHSA	Q6R1D6	sars	corona
612	15	46.9	99	2	Q71267	9H1V1	Q71267	human	immun	685	15	46.9	122	2	Q76XZ9	CVHSA	Q76XZ9	sars	corona
613	15	46.9	99	2	Q90DI1	9H1V1	Q90DI1	human	immun	686	15	46.9	122	2	Q6RCV9	CVHSA	Q6RCV9	sars	corona
614	15	46.9	99	2	Q90DI2	9H1V1	Q90DI2	human	immun	687	15	46.9	122	2	Q6RCX0	CVHSA	Q6RCX0	sars	corona
615	15	46.9	100	2	Q90DI5	9H1V1	Q90DI5	human	immun	688	15	46.9	122	2	Q6RCX0	CVHSA	Q6RCX0	sars	corona

689	15	46.9	122	2	Q6RCY1_CVHSA	Q6rcy1	sars	corona	762	15	46.9	136	2	Q4V6K1_DROME	Q4v6k1	drosophila
690	15	46.9	122	2	Q6RCZ2_CVHSA	Q6rcz2	sars	corona	763	15	46.9	136	2	Q6GRP1_ORYSA	Q6grp1	oryza sativ
691	15	46.9	122	2	Q6RD03_CVHSA	Q6rd03	sars	corona	764	15	46.9	137	2	Q6X1U8_DROYA	Q6x1u8	drosophila
692	15	46.9	122	2	Q6RD14_CVHSA	Q6rd14	sars	corona	765	15	46.9	137	2	Q63U00_BURPS	Q63u00	burkholderi
693	15	46.9	122	2	Q6RD25_CVHSA	Q6rd25	sars	corona	766	15	46.9	137	2	Q5FXE1_9HEPC	Q5fxe1	hepatitis c
694	15	46.9	122	2	Q6RD36_CVHSA	Q6rd36	sars	corona	767	15	46.9	137	2	Q5FXE2_9HEPC	Q5fxe2	hepatitis c
695	15	46.9	122	2	Q6RD47_CVHSA	Q6rd47	sars	corona	768	15	46.9	137	2	Q5FXE2_9HEPC	Q5fxe2	hepatitis c
696	15	46.9	122	2	Q6RD58_CVHSA	Q6rd58	sars	corona	769	15	46.9	137	2	Q5FXE3_9HEPC	Q5fxe3	hepatitis c
697	15	46.9	122	2	Q6RD65_CVHSA	Q6rd65	sars	corona	770	15	46.9	137	2	Q5FXG3_9HEPC	Q5fxg3	hepatitis c
698	15	46.9	122	2	Q6RR05_CVHSA	Q6rr05	sars	corona	771	15	46.9	137	2	Q5FXG9_9HEPC	Q5fxg9	hepatitis c
699	15	46.9	122	2	Q6RRR5_CVHSA	Q6rrr5	sars	corona	772	15	46.9	137	2	Q5FXH0_9HEPC	Q5fxh0	hepatitis c
700	15	46.9	122	2	Q6RRS0_CVHSA	Q6rrs0	sars	corona	773	15	46.9	137	2	Q5FXH1_9HEPC	Q5fxh1	hepatitis c
701	15	46.9	122	2	Q6RRS5_CVHSA	Q6rrs5	sars	corona	774	15	46.9	137	2	Q5FXH2_9HEPC	Q5fxh2	hepatitis c
702	15	46.9	122	2	Q6SRT0_CVHSA	Q6srt0	sars	corona	775	15	46.9	137	2	Q5FXH3_9HEPC	Q5fxh3	hepatitis c
703	15	46.9	122	2	Q6SRT5_CVHSA	Q6srt5	sars	corona	776	15	46.9	137	2	Q5FXH4_9HEPC	Q5fxh4	hepatitis c
704	15	46.9	122	2	Q6SRU0_CVHSA	Q6sru0	sars	corona	777	15	46.9	137	2	Q5FXH6_9HEPC	Q5fxh6	hepatitis c
705	15	46.9	122	2	Q6SRU5_CVHSA	Q6sru5	sars	corona	778	15	46.9	137	2	Q5FXH7_9HEPC	Q5fxh7	hepatitis c
706	15	46.9	122	2	Q6SRV0_CVHSA	Q6srv0	sars	corona	779	15	46.9	137	2	Q5FXH7_9HEPC	Q5fxh7	hepatitis c
707	15	46.9	122	2	Q6SRV5_CVHSA	Q6srv5	sars	corona	780	15	46.9	137	2	Q4S169_TETNG	Q4s169	tetradon n
708	15	46.9	122	2	Q6SRW0_CVHSA	Q6srw0	sars	corona	781	15	46.9	138	2	Q5Y9V8_LYCCE	Q5y9v8	lycopersico
709	15	46.9	122	2	Q6SRW5_CVHSA	Q6srw5	sars	corona	782	15	46.9	138	2	Q7FEB7_RHCOM	Q7feb7	rhebus cyto
710	15	46.9	122	2	Q6SRX0_CVHSA	Q6srx0	sars	corona	783	15	46.9	139	2	Q5GY53_XANOR	Q5gy53	xanthomonas
711	15	46.9	122	2	Q6SRX5_CVHSA	Q6srx5	sars	corona	784	15	46.9	139	2	Q4MNG4_BACCE	Q4mng4	bacillus ce
712	15	46.9	122	2	Q6SRV0_CVHSA	Q6sry0	sars	corona	785	15	46.9	139	2	Q6HFI1_BACHK	Q6hfi1	bacillus th
713	15	46.9	122	2	Q6SRV5_CVHSA	Q6sry5	sars	corona	786	15	46.9	139	2	Q733N2_BACCL	Q733n2	bacillus ce
714	15	46.9	122	2	Q6SRZ0_CVHSA	Q6srz0	sars	corona	787	15	46.9	139	2	Q81AG4_BACCR	Q81ag4	bacillus ce
715	15	46.9	122	2	Q6VA72_CVHSA	Q6va72	sars	corona	788	15	46.9	139	2	Q97D28_CIOAB	Q97d28	clostridium
716	15	46.9	122	2	Q6VA83_CVHSA	Q6va83	sars	corona	789	15	46.9	139	2	Q81Y90_BACAN	Q81y90	bacillus an
717	15	46.9	122	2	Q6VA94_CVHSA	Q6va94	sars	corona	790	15	46.9	139	2	Q637L9_BACCZ	Q637l9	bacillus ce
718	15	46.9	122	2	Q4JDPE_CVHSA	Q4jdp8	sars	corona	791	15	46.9	139	2	Q69963_9H1V1	Q69963	human immun
719	15	46.9	122	2	Q4JDRI_CVHSA	Q4jdr1	sars	corona	792	15	46.9	140	1	LY6H_HUMAN	LY6H	HUMAN
720	15	46.9	122	2	Q6C049_XENLA	Q6c049	xenopus	lae	793	15	46.9	140	2	Q6TAX0_HUMAN	Q6tax0	homo sapien
721	15	46.9	123	2	Q9JQZ2_NEITMA	Q9jqz2	neisseria	m	794	15	46.9	140	2	Q4V3U1_DROME	Q4v3u1	drosophila
722	15	46.9	123	2	Q7DDH5_NEITMB	Q7ddh5	neisseria	m	795	15	46.9	140	2	Q8GZ25_ORYSA	Q8gz25	oryza sativ
723	15	46.9	123	2	Q6Y908_9H1V1	Q6y908	human	immun	796	15	46.9	140	2	Q8AR11_9H1V1	Q8ar11	human immun
724	15	46.9	124	2	Q9PKS3_CHLMU	Q9pks3	chlamydia	m	797	15	46.9	141	2	Q6GKZ6_DROME	Q6gkz6	drosophila
725	15	46.9	124	2	Q7SR41_9H1V1	Q7sr41	human	immun	798	15	46.9	141	2	Q664U8_YERPS	Q664u8	yersinia ps
726	15	46.9	125	2	Q6KD02_ECOLI	Q6kd02	escherichia		799	15	46.9	141	2	Q8ZJ85_YERPE	Q8zj85	yersinia pe
727	15	46.9	125	2	Q4K3F0_PSEF5	Q4k3f0	pseudomonas		800	15	46.9	142	2	Q9VF15_DROME	Q9vf15	drosophila
728	15	46.9	125	2	Q8FI23_ECOL6	Q8fi23	escherichia		801	15	46.9	143	2	Q6ZTJ0_HUMAN	Q6ztj0	homo sapien
729	15	46.9	125	2	Q4TEY5_TETNG	Q4tey5	tetradon n		802	15	46.9	143	2	Q8S406_HORVD	Q8s406	hordewum vul
730	15	46.9	125	2	Q6RUW8_9H1V1	Q6ruw8	human	immun	803	15	46.9	143	2	Q9ACU2_STRCO	Q9acu2	strepococ
731	15	46.9	126	2	Q4I7X6_GIBZE	Q4i7x6	gibberella		804	15	46.9	143	2	Q6WH33_BDEBA	Q6wh33	belioliobri
732	15	46.9	126	2	Q7R635_GIALA	Q7r635	giardia lam		805	15	46.9	144	2	Q4F8J1_EUCGI	Q4f8j1	eucalyptus
733	15	46.9	126	2	Q88C21_PSEPK	Q88c21	pseudomonas		806	15	46.9	144	2	Q6CZ18_YERPE	Q6cz18	yersinia pe
734	15	46.9	127	2	Q4ZL69_PSEBSY	Q4zl69	pseudomonas		807	15	46.9	145	2	Q8WQ22_LOOMI	Q8wq22	loocusta mlg
735	15	46.9	127	2	Q69962_9H1V1	Q69962	human	immun	808	15	46.9	145	2	Q9VZT5_DROME	Q9vzt5	drosophila
736	15	46.9	128	1	KRA24_HUMAN	Q9byr9	homo sapien		809	15	46.9	145	2	Q6Z626_ORYSA	Q6z626	oryza sativ
737	15	46.9	128	1	Q8LBO3_ARATH	Q8lbg3	arabidopsis		810	15	46.9	146	2	Q5KAF7_SCHGR	Q5kaf7	schistocerc
738	15	46.9	128	2	Q4ZU08_PSEBSY	Q4zu08	pseudomonas		811	15	46.9	146	2	Q7XNZ5_ORYSA	Q7xnz5	oryza sativ
739	15	46.9	129	2	Q5D9M3_SCHJA	Q5dgm3	schistosoma		812	15	46.9	146	2	Q8AXW7_MIGCO	Q8axw7	micrurus co
740	15	46.9	129	2	Q6IKR4_DROME	Q6ikr4	drosophila		813	15	46.9	146	2	Q4REX5_TETNG	Q4rex5	tetradon n
741	15	46.9	129	2	Q50KR5_LYVMT	Q50kr5	lymaea ste		814	15	46.9	147	2	Q5Y2U3_9H1V1	Q5y2u3	human immun
742	15	46.9	129	2	Q4IYL5_AZQVI	Q4iyl5	azotobacter		815	15	46.9	148	2	Q6ZW92_HUMAN	Q6zw92	homo sapien
743	15	46.9	130	2	Q5AOR7_DICDI	Q5aor7	dictyosteli		816	15	46.9	148	2	Q5TC71_HUMAN	Q5tc71	homo sapien
744	15	46.9	130	2	Q6THB0_DROME	Q6thb0	drosophila		817	15	46.9	148	2	Q4RL15_TETNG	Q4rl15	tetradon n
745	15	46.9	131	2	Q9S237_STRCO	Q9s237	streptomyce		818	15	46.9	149	2	Q8UAP1_AGRTS	Q8uap1	agrobacteri
746	15	46.9	131	2	Q6CZ68_ERWCT	Q6cz68	erynia car		819	15	46.9	149	2	Q898V0_CIOTE	Q898v0	clostridium
747	15	46.9	132	2	Q55ZH5_CRYNE	Q55zh5	cryptococcu		820	15	46.9	150	2	Q5AC28_CANAL	Q5ac28	candida alb
748	15	46.9	132	2	Q5D995_SCHJA	Q5d995	schistosoma		821	15	46.9	150	2	Q6ICP4_CABBR	Q6icp4	caenorhabdi
749	15	46.9	132	2	Q9AN59_BRATJ	Q9an59	bradyrhizob		822	15	46.9	150	2	Q6LDD5_PARTE	Q6ldd5	paramechiu
750	15	46.9	132	2	Q7SOV8_9H1V1	Q7sov8	human	immun	823	15	46.9	151	2	Q9BMH5_CABEL	Q9bmh5	caenorhabdi
751	15	46.9	133	2	Q4T2U4_TETNG	Q4t2u4	tetradon n		824	15	46.9	152	2	Q8TIR1_METAC	Q8tir1	methanobac
752	15	46.9	133	2	Q8UUD5_9H1V1	Q8uud5	human	immun	825	15	46.9	152	2	Q7PX28_ANGOA	Q7px28	anopheles g
753	15	46.9	134	2	Q24959_GIALA	Q24959	giardia lam		826	15	46.9	153	2	Q80876_BSMY	Q80876	barley str1
754	15	46.9	134	2	Q24988_GIALA	Q24988	giardia lam		827	15	46.9	153	2	Q94L78_ORYSA	Q94l78	oryza sativ
755	15	46.9	134	2	Q4V6M8_DROME	Q4v6m8	drosophila		828	15	46.9	154	2	Q5TRK5_ANGOA	Q5trk5	anopheles g
756	15	46.9	134	2	Q6RUX2_9H1V1	Q6rux2	human	immun	829	15	46.9	155	2	Q95KX0_BOVIN	Q95kx0	bos taurus
757	15	46.9	135	1	YDQA_SCHPO	Q14202	schizosacch		830	15	46.9	155	2	Q8GSR9_WHBAT	Q8gsr9	tritricum ae
758	15	46.9	135	2	Q5D8E6_SCHJA	Q5d8e6	schistosoma		831	15	46.9	155	2	Q69536_9SBTA	Q69536	human hepe
759	15	46.9	135	2	Q4QIR1_LEHMA	Q4qir1	leishmania		832	15	46.9	155	2	Q12317_9H1V1	Q12317	human immun
760	15	46.9	135	2	Q84YV6_ORYSA	Q84yv6	oryza sativ		833	15	46.9	156	2	Q8JMH4_9NUCL	Q8jmh4	manebstra co
761	15	46.9	135	2	Q8H6Y5_PHTIN	Q8h6y5	phytophthor		834	15	46.9	157	1	CI025_HUMAN	CI025	homo sapien



```

981 15 46.9 207 2 04V6B8 DROME 04V6B8 drosophila
982 15 46.9 207 2 070016 SHIDY 070016 shigella dy
983 15 46.9 207 2 08FCJ3 ECOL6 08FCJ3 escherichia
984 15 46.9 207 2 08XSN4 EC057 08XSN4 escherichia
985 15 46.9 208 2 06XHV9 DROXYA 06XHV9 drosophila
986 15 46.9 208 2 082069 SOLTU 082069 solanum tub
987 15 46.9 208 2 04PBJ3 EUGLU 04PBJ3 eucalyptus
988 15 46.9 208 2 09AHJ9 GLACO 09AHJ9 lactobacilli
989 15 46.9 209 2 05SGM2 DICTI 05SGM2 dictyosteli
990 15 46.9 210 2 07KML7 DICTI 07KML7 dictyosteli
991 15 46.9 210 2 07XLI7 ORYSA 07XLI7 oryza sativ
992 15 46.9 210 2 093CB9 MYCPA 093CB9 mycobacteri
993 15 46.9 210 2 039318 PHTV1 039318 human immun
994 15 46.9 211 1 G12 ANOGA 017040 anopheles g
995 15 46.9 211 2 07Q5V5 ANOGA 07Q5V5 anopheles g
996 15 46.9 211 2 050SC9 ENTHI 050SC9 entamoeba h
997 15 46.9 211 2 06WES3 PTILE 06WES3 polliachius
998 15 46.9 212 2 050PES ENTHI 050PES entamoeba h
999 15 46.9 212 2 09XV06 CAEBL 09XV06 caenorhabdi
1000 15 46.9 212 2 04Q6X5 LEIMA 04Q6X5 leishmania

```

## ALIGNMENTS

```

RESULT 1
096KM3_HUMAN PRELIMINARY; PRT; 213 AA.

```

```

ID 096KM3_HUMAN PRELIMINARY; PRT; 213 AA.
AC 096KM3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Aberrant WW domain-containing oxidoreductase.
GN Name=WWOX;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxId=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21457335; PubMed=11572989; DOI=10.1073/pnas.191175896;
RA Paige A.J.W., Taylor K.J., Taylor C., Hillier S.G., Farrington S.,
RA Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.;
RT "WWOX: a candidate tumor suppressor gene involved in multiple tumor
RT types.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422(2001).
RL EMBL; AF325432; AAL05451.1; -; Genomic DNA.
DR EMBL; AF325423; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325424; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325426; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325433; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325425; AAL05451.1; JOINED; Genomic DNA.
DR HSSP; Q13526; IPIN.
DR Ensembl; ENSG00000186153; Homo sapiens.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS00159; WW_DOMAIN_1; 2.
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
SQ SEQUENCE 213 AA; 23866 MW; A21054FF8214C7C CRC64;

```

```

Query Match 56.2%; Score 18; DB 2; Length 213;
Best Local Similarity 20.0%; Pred. No. 0.43;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 4 CXXXXXXXXX 13
DB 184 CTSTRAAAC 193

```

```

RESULT 2
Q586F4_9TRYP

```

```

ID 0586F4_9TRYP PRELIMINARY; PRT; 263 AA.
AC 0586F4_9TRYP;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=TP927.6.4820;
OS Trypanosoma brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxId=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Shadlow E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA Khalak H.G., Salberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wanstall D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA El-Sayed N.M., Khalak H., Adams M.D.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Haas B., Blandin G., El-Sayed N.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC008146; AAX80300.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 263 AA; 28873 MW; F61DF5AF910531AB CRC64;

```

```

Query Match 56.2%; Score 18; DB 2; Length 263;
Best Local Similarity 20.0%; Pred. No. 0.46;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 4 CXXXXXXXXX 13
DB 95 CATTSSSAC 104

```

```

RESULT 3
07PRO7_ANOGA PRELIMINARY; PRT; 269 AA.
ID 07PRO7_ANOGA PRELIMINARY; PRT; 269 AA.
AC 07PRO7;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ENSANGP0000001657 (Fragment).
GN ORFNames=ENSANG0000001387;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxId=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008847; EAA06779.3; -; Genomic DNA.
FT NON_TER 1 1
FT 269 269

```

SQ SEQUENCE 269 AA; 24228 MW; 25BBFF71FD71F1F2 CRC64;  
 Query Match 56.2%; Score 18; DB 2; Length 269;  
 Best Local Similarity 20.0%; Pred. No. 0.47;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 DB 149 CTSSTTSC 158

RESULT 4  
 05TX10 ANOGA PRELIMINARY; PRT; 281 AA.  
 ID 05TX10;  
 AC 01-FEB-2005 (T-EMBLrel. 29, Created)  
 DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)  
 DE ENSANGP0000027669 (Fragment).  
 GN ORFNames=ENSANG0000007023;  
 OS Anopheles gambiae str. PE8T.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;  
 OC Anophelinae; Anopheles.  
 NCBI\_TaxID=180454;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=PE8T.  
 RG The Anopheles gambiae Sequence Committee;  
 RT "Anopheles gambiae re-annotation."  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=PE8T.  
 RG The Anopheles gambiae Sequence Committee;  
 RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAAAB01008807; EAL41760.1; -; Genomic\_DNA.  
 FT NON\_TER 1  
 FT NON\_TER 281  
 SQ SEQUENCE 281 AA; 25341 MW; 0462E1169FB8642B CRC64;

Query Match 56.2%; Score 18; DB 2; Length 281;  
 Best Local Similarity 20.0%; Pred. No. 0.48;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 DB 107 CTAASSTAC 116

RESULT 5  
 YMS8 YEAST STANDARD; PRT; 313 AA.  
 ID YMS8 YEAST  
 AC 003635;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Hypothetical 35.0 kDa protein in PRK2-HF1 intergenic region.  
 GN OrderedAccession=YMR206W; ORFNames=YMR325.07;  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NCBI\_TaxID=4932;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=S288c / AB972;  
 RX MEDLINE=97313268; PubMed=9169872;  
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,  
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,  
 RA Jagers K., Lye G., Moule S., Odeli C., Pearson D., Rajandream M.A.,

RA Rice P., Skelton J., Walsh S.V., Whitehead S., Barrell B.G.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RT XIII."  
 RL Nature 387:90-93 (1997).

CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL; 248755; CAA88648.1; -; Genomic\_DNA.  
 DR PIR; S59448; S59448.  
 DR GenomOnline; 142881; -;  
 DR Ensembl; YMR206W; Saccharomyces cerevisiae.  
 DR SCD; S000004819; YMR206W.  
 KW Complete proteome; Hypothetical protein.  
 FT COMPBIAS 3  
 FT COMPBIAS 146 149 Poly-Ser.  
 FT COMPBIAS 246 252 Poly-Gln.  
 FT COMPBIAS 246 252 Poly-Ser.  
 SQ SEQUENCE 313 AA; 35018 MW; 9D92BFD982577F0 CRC64;

Query Match 56.2%; Score 18; DB 1; Length 313;  
 Best Local Similarity 20.0%; Pred. No. 0.5;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 DB 245 CSSSSSSSNC 254

RESULT 6  
 Q6ZDR4 ORYSA PRELIMINARY; PRT; 357 AA.  
 ID Q6ZDR4 ORYSA  
 AC Q6ZDR4;  
 DT 05-JUL-2004 (T-EMBLrel. 27, Created)  
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)  
 DE Putative transcription factor Myb protein.  
 GN Name=P0481F05.20;  
 OS Oryza sativa (japonica cultivar-group);  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoideae; Oryza.  
 NCBI\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC  
 RT clone: P0481F05."  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 DR EMBL; AP004376; BAD09322.1; -; Genomic\_DNA.  
 DR HSSP; O03237; IASJ.  
 DR Gramine; Q6ZDR4; -;  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR InterPro; IPR001005; Myb\_DNA\_binding; IEA.  
 DR InterPro; IPR012287; Homeodomain-rel.  
 DR Pfam; PF00249; Myb\_DNA-binding; 2.  
 DR SMART; SM00717; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; UNKNOWN\_1.  
 DR PROSITE; PS00034; MYB\_1; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KW Nuclear protein; Repeat.  
 SQ SEQUENCE 357 AA; 36754 MW; 3F4F56860D3B4E22 CRC64;

Query Match 56.2%; Score 18; DB 2; Length 357;  
 Best Local Similarity 20.0%; Pred. No. 0.52;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

Db 322 CSATASASC 331

RESULT 7

Q7QJ10 ANOGA PRELIMINARY; PRT; 362 AA.

ID Q7QJ10.0

AC Q7QJ10.0

DT 01-MAR-2004 (TRENBLREL. 26, Created)

DT 01-MAR-2004 (TRENBLREL. 26, Last sequence update)

DE ENSANGP000009383 (Fragment).

GN ORFNames=ENSANG0000007023;

OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae; Anophelinae; Anopheles.

OC NCBI\_TaxID=180454;

OK NCBI\_TaxID=180454;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RT "Anopheles gambiae re-annotation."

RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.

CC -! CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

DR EMBL; AAAB01008807; EAA04378.2; -; Genomic\_DNA.

FT NON\_TER 1

SQ SEQUENCE 362 AA; 32490 MW; 384376DFA31BEC0 CRC64;

Query Match 56.2%; Score 18; DB 2; Length 362;

Best Local Similarity 20.0%; Pred. No. 0.52;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

Db 187 CTAASSTAC 196

RESULT 8

Q9RV87 DEIRA PRELIMINARY; PRT; 480 AA.

ID Q9RV87

AC Q9RV87

DT 01-MAY-2000 (TRENBLREL. 13, Created)

DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)

DE Hypothetical protein DR1142.

GN OrderedLOCNames=DR1142.

OS Deinococcus radiodurans.

OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales; Deinococcaceae; Deinococcus.

OC NCBI\_TaxID=1299;

OK NCBI\_TaxID=1299;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;

RX MEDLINE=2003696; PubMed=10567266; DOI=10.1126/science.286.5444.1571;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L., Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M., Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D., Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C., Frazer C.M.;

RT "Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1."

RL Science 286:1571-1577(1999).

DR EMBL; AE001963; AAF10716.1; -; Genomic\_DNA.

PIR; E75433; E75433.

DR TIGR; DR1142; -;

KM Complete proteome; Hypothetical protein.

SQ SEQUENCE 480 AA; 49364 MW; 18B962AD5A24AE CRC64;

Query Match 56.2%; Score 18; DB 2; Length 480;

Best Local Similarity 20.0%; Pred. No. 0.58;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

Db 464 C5AATAATAC 473

RESULT 9

TMCC1 HUMAN STANDARD; PRT; 653 AA.

ID TMCC1.HUMAN

AC 094876; Q68E06; Q8IXM8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 10-MAY-2005 (Rel. 47, Last sequence update)

DE Transmembrane and coiled-coil domain protein 1.

GN Name=TMCC1; Synonyms=K1AA0779;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

OK NCBI\_TaxID=9606;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RC TISSUE=Skin;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uddin T.B., Tashy S., Carninci P., Prange C., Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullihy S.J., Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Gardia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.B., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 15-653.

RC TISSUE=Endometrial tumor;

RG The German cDNA consortium;

RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 334-653.

RC TISSUE=Brain;

RX MEDLINE=99087487; PubMed=9872452;

RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."

RL DNA Res. 5:277-286(1998).

CC -! SIMILARITY: Belongs to the TEX28 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

```

CC -----
DR EMBL; AB018322; BAA34499.1; -; mRNA.
DR EMBL; BC039859; AAH39859.1; -; mRNA.
DR EMBL; CR749206; CAH18064.2; -; mRNA.
DR Ensembl; ENSG00000172765; Homo sapiens.
DR HGN; HGNC:29116; TMCC1.
KW Coiled coil; Transmembrane.
FT TRANSMEM 592 612 Potential.
FT TRANSMEM 625 645 Potential.
FT COILED 228 313 Potential.
FT COILED 458 576 Potential.
SQ SEQUENCE 653 AA; 72053 MW; 18B07D11B874205 CRC64;

Query Match
Best Local Similarity 56.2%; Score 18; DB 1; Length 653;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 173 CAAAAAAAC 182

RESULT 10
Q5317_XANOR PRELIMINARY; PRT; 1005 AA.
ID Q5317_XANOR PRELIMINARY; PRT; 1005 AA.
AC Q5H317;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Transcriptional regulator.
DE Name:acok; OrderedlocusNames=XO01230;
OS Xanthomonas oryzae (pv. oryzae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OC NCBI_TaxId=64187;
RN (1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KACC10331 / KX085;
RX PubMed=15673718; DOI=10.1093/nar/gki206;
RA Lee B.-M., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
RA Park I.-C., Yoon U.-H., Hahn J.-H., Koo B.-S., Lee G.-B., Kim H.,
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
RA Go S.-J.;
RT "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,
RT the bacterial blight pathogen of rice.";
RL Nucleic Acids Res. 33:577-586(2005).
DR EMBL; AE013598; AAU74484.1; -; Genomic DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:001111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR000792; HTH_LuxR.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR011717; TPR_4.
DR Pfam; PF00196; Gere; 1.
DR Pfam; PF00515; TPR_1; 2.
DR Pfam; PF07721; TPR_4; 1.
DR PRINTS; PR00038; HTHLUXR.
DR ProDom; PD000307; HTH_LuxR; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00421; HTH_LuxR; 1.
DR SMART; SM00026; TPR; 3.
DR PROSITE; PS00622; HTH_LuxR_FAMILY; 1.
DR PROSITE; PS00293; TPR_REGION; 1.
DR PROSITE; PS00293; TPR_REGION; 1.
KW Complete proteome.
SQ SEQUENCE 1005 AA; 110791 MW; 662961379FD37B61 CRC64;

Query Match
Best Local Similarity 56.2%; Score 18; DB 2; Length 1005;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 4 CXXXXXXXXX 13
DB 386 CASATTAAC 395

RESULT 11
Q4QCK7_LEIMA PRELIMINARY; PRT; 1061 AA.
ID Q4QCK7_LEIMA PRELIMINARY; PRT; 1061 AA.
AC Q4QCK7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Kinesin heavy chain, putative.
DE ORFNames=ImjF20.0640;
OS Leishmania major.
OC Eukaryota; Eukaryozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OC NCBI_TaxId=5664;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005259; CAJ03791.1; -; Genomic DNA.
SQ SEQUENCE 1061 AA; 113930 MW; 2C69385D06371DCD CRC64;

Query Match
Best Local Similarity 56.2%; Score 18; DB 2; Length 1061;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 192 CAATATSTSC 201

RESULT 12
Q8LPD6_HORVU PRELIMINARY; PRT; 52 AA.
ID Q8LPD6_HORVU PRELIMINARY; PRT; 52 AA.
AC Q8LPD6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative glutathione-S-transferase (Fragment).
DE Name=bar2;
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OC NCBI_TaxId=4513;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RA Baldwin A.;
RT "The effect of thiocarbamate herbicides and the safener dichlormid in
RT barley and wild cats.";
RL Thesis (2001), Department of Cardiff School of Biosciences, Cardiff
RL university, Cardiff, United Kingdom.
RL [2]
RP NUCLEOTIDE SEQUENCE.
RA Baldwin A., Harwood J.L., Machray G.C., Francis D., Rogers H.J.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ419772; CAD11963.1; -; mRNA.
DR HSSP; P12653; IAXD.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002197; HTH_Fis.
DR PRINTS; PR01590; HTHFIS.
KW Transferase.
FT NON_TER 1 1
SQ SEQUENCE 52 AA; 5750 MW; 7392EB72A3C70B9B CRC64;

```



Query Match 53.1%; Score 17; DB 2; Length 52;  
 Best Local Similarity 20.0%; Pred. No. 1.3;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 13  
 DB 18 CTSASSTRSC 27

## RESULT 13

OSDCN3 SCHJA PRELIMINARY; PRT; 122 AA.  
 AC QSDCN3;  
 DT 10-MAY-2005 (TReMBLrel. 30, Created)  
 DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)  
 DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)  
 DE Hypothetical protein.  
 OS Schistosoma japonicum (Blood fluke).  
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidae;  
 OC Schistosomatidae; Schistosomatidae; Schistosoma.  
 OX NCBI\_TaxID=6182;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Han Z.;  
 RT "The full-length cDNA sequences of Schistosoma japonicum genes."  
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY814691; AAW26423.1; -; mRNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 122 AA; 13028 MW; 960B8615986958DE CRC64;

Query Match 53.1%; Score 17; DB 2; Length 122;  
 Best Local Similarity 20.0%; Pred. No. 1.8;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 13  
 DB 76 CTSSTSTSC 85

## RESULT 14

O96LJ4 HUMAN PRELIMINARY; PRT; 128 AA.  
 AC O96LJ4;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Hypothetical protein FLJ25437.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Testis;  
 RA Kawakami B., Sugiyama A., Takemoto M., Suzuki Y., Hata H.,  
 RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,  
 RA Irie R., Otsuki T., Sato H., Nishikawa T., Nagai K., Isogai T.,  
 RA Sugano S.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK058166; BAB71697.1; -; mRNA.  
 DR Ensembl; ENSG00000178093; Homo sapiens.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004672; F:protein kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR ATP-binding; Kinase; Nucleotide-binding; Transferase;  
 SQ SEQUENCE 128 AA; 14116 MW; D30AC8C668621B0C CRC64;

Query Match 53.1%; Score 17; DB 2; Length 128;  
 Best Local Similarity 20.0%; Pred. No. 1.8;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 13  
 DB 118 CTAISSAKTC 127

## RESULT 15

O6ZKM7 ORYSA PRELIMINARY; PRT; 139 AA.  
 AC O6ZKM7;  
 DT 05-JUL-2004 (TReMBLrel. 27, Created)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE Hypothetical protein OJ118\_A06.7-1.  
 GN Name=OJ118\_A06.7-1;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsunoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, BAC  
 clone:OJ118\_A06.7-1."  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003873; BAD08805.1; -; Genomic DNA.  
 KW Gramene; O6ZKM7;  
 KM Hypothetical protein.  
 SQ SEQUENCE 139 AA; 14109 MW; E281360370FB63C9 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 139;  
 Best Local Similarity 20.0%; Pred. No. 1.9;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 13  
 DB 77 CAITSSSSC 86

## RESULT 16

O6H482 ORYSA PRELIMINARY; PRT; 155 AA.  
 AC O6H482;  
 DT 05-JUL-2004 (TReMBLrel. 27, Created)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
 DT 10-MAY-2005 (TReMBLrel. 30, Last annotation update)  
 DE Hypothetical protein OJ119\_C03.6 (Hypothetical protein  
 DE B1012G11.50).  
 GN Name=OJ119\_C03.6; Synonyms=B1012G11.50;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsunoto T., Hattori M., Sasaki Y., Katayose Y.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, BAC  
 clone:OJ119\_C03.6."  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsunoto T., Katayose Y.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, BAC  
 clone:B101G11.1."  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP005969; BAD26467.1; -; Genomic DNA.  
 DR EMBL; AP005903; BAD38424.1; -; Genomic DNA.



DR Gramene; Q6H482; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 155 AA; 16233 MW; 02D8365B8C7B52 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 155;  
 Best Local Similarity 20.0%; Pred. No. 2;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 DB 25 CSTSAADC 34

RESULT 17  
 O657M0\_ORYSA PRELIMINARY; PRT; 160 AA.

AC O657M0; 25-OCT-2004 (TRENBLrel. 28, Created)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)  
 DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)  
 DE Hypothetical protein P0468H06.13 (Hypothetical protein  
 P0468H05.38).  
 GN Name=P0468H06.13; Synonyms=P0468H05.38;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoideae; Oryza.  
 NC NCBT\_TaxID=39947;  
 RN (1)

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Maksumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
 RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
 RA Ikemoto M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,  
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
 RA Shinokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshitani R., Yukawa K.,  
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.;  
 RT "The genome sequence and structure of rice chromosome 1.";  
 RL Nature 420:312-316(2002).  
 DR EMBL; AP003075; BAD44997.1; -; Genomic DNA.  
 DR EMBL; AP003267; BAD87274.1; -; Genomic DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 160 AA; 17630 MW; C3A5AED5C75DA9 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 160;  
 Best Local Similarity 20.0%; Pred. No. 2;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 DB 105 CSGATSTASC 114

RESULT 18  
 O4T4H1\_TETNG PRELIMINARY; PRT; 160 AA.

AC O4T4H1; 13-SEP-2005 (TRENBLrel. 31, Created)  
 DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)  
 DE Chromosome undetermined SCAP9659, whole genome shotgun sequence.  
 GN ORFName=GSTENG00007226001.  
 OS Tetradodon nigrovittatus (Green puffer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Perciformes; Tetradontiformes;  
 OC Tetradontidae; Tetradontidae; Tetradon.  
 NC NCBT\_TaxID=99883;  
 RN (1)

RP NUCLEOTIDE SEQUENCE.

RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,  
 RA Mauceli E., Bouneau L., Fischer C., Oudet-Costat C., Bernot A.,  
 RA Micaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segreus B.,  
 RA Dasila C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,  
 RA Blemont C., Skalli Z., Cateolico L., Poulain J., De Bernardis V.,  
 RA Cruaud C., Duprat S., Brothier P., Contancieu J.P., Gouzy J.,  
 RA Parra G., Lardier G., Chapelle C., McKernan K.J., McGowan P., Bosak S.,  
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,  
 RA Wincker P., Lander E.S., Weissbach J., Roest Crollins H.,  
 RT "Genome duplication in the teleost fish Tetradodon nigrovittatus reveals  
 the early vertebrate proto-karyotype.";  
 RL Nature 431:946-957(2004).  
 RN (2)

RP NUCLEOTIDE SEQUENCE.

RG Genoscope, Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.

DR EMBL; CAAB0100659; CAF9221.1; -; Genomic DNA.  
 SQ SEQUENCE 160 AA; 17094 MW; 61673E4F2AE5451B CRC64;

Query Match 53.1%; Score 17; DB 2; Length 160;  
 Best Local Similarity 20.0%; Pred. No. 2;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 DB 115 CTTARSSTSC 124

RESULT 19  
 O825B8\_STRAW PRELIMINARY; PRT; 219 AA.

AC O825B8; 01-JUN-2003 (TRENBLrel. 24, Created)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Putative tetr-family transcriptional regulator.  
 GN OrderedListNames=SAV7510;  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomyces; Streptomyces; Streptomyces.  
 NC NCBT\_TaxID=33903;  
 RN (1)

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RA STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nb820;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shida T.,  
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 RN (2)

RP NUCLEOTIDE SEQUENCE.

RA STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nb820;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shida T.,  
 RA Sakaki Y., Hattori M., Omura S.;  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis.";  
 RL Nat. Biotechnol. 21:526-531(2003).  
 DR EMBL; BA000030; BAC75221.1; -; Genomic DNA.

```
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR012287; Homodomain-rel.
DR InterPro; IPR01647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PROSITE; PS50977; HTH_TERR_2; 1.
KM Complete proteome.
SQ SEQUENCE 219 AA; 24066 MW; 5C59ADF39430F304 CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 219;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 151 CAEAAAAAAC 160

RESULT 20
Q6152_YVIRU PRELIMINARY; PRT; 220 AA.
ID Q6152_YVIRU
AC Q6152;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Coat protein.
OS Cymbidium mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flexiviridae;
OC Potexvirus.
OX NCBI_TaxID=12178;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Korean isolate;
RA MEDLINE:95278762; PubMed:7758973; DOI=10.1016/0378-1119(95)00105-F;
RA Ryu K.H., Yoon K.B., Park W.M.;
RT "Nucleotide sequence of coat protein gene of cymbidium mosaic
RT potexvirus genomic RNA, the Korean isolate.";
RL Gene 156:303-304(1995).
DR EMBL; X81051; CAA56941.1; -; Genomic_RNA.
DR PIR; JC4082; JC4082.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000052; P1vtr_coat.
DR Pfam; PF00286; Virus_P-coat; 1.
DR PRINTS; PR00232; POTXCARLCOAT.
DR ProDom; PD00603; P1vtr_coat; 1.
KM Capsid protein.
SQ SEQUENCE 220 AA; 23719 MW; 1F40C6F7C0CCF3B8 CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 220;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 166 CSAATLTATC 175

RESULT 21
Q6YTS2_ORYSA PRELIMINARY; PRT; 256 AA.
ID Q6YTS2_ORYSA
AC Q6YTS2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein P0419H09.29.
GN Name=P0419H09.29;
OS Oryza sativa (Japonica cultivar-group). Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RA Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005918; BAD10719.1; -; Genomic_DNA.
DR Gramene; O6YTS2; -.
KM Hypothetical protein.
SQ SEQUENCE 256 AA; 26591 MW; 63BFCAF3D2D40EA CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 256;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 203 CSASASARC 212

RESULT 22
Q850Z3_ORYSA PRELIMINARY; PRT; 360 AA.
ID Q850Z3_ORYSA
AC Q850Z3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE DnaJ-like protein (putative heat shock protein).
GN Name=OSUNBD00277B08.12; Synonym=OSUNBa0078D06.16;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buehl C.R., Yuan Q., Ouyang S., Liu J., Ganabergner K., Jones K.M.,
RA Overell II L.L., Tselirin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmiller S.B., Utehrack T.T., Feldlyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackebush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buehl R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Buehl R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Have a continuous role in plant development probably in
CC the structural organization of compartments (by similarity).
DR EMBL; AC133778; AA039854.1; -; Genomic_DNA.
DR EMBL; AC133339; AAP46230.1; -; Genomic_DNA.
DR HSSP; P25685; 1HDJ.
DR Gramene; Q850Z3; -.
DR GO; GO:0031072; F:heat shock protein binding; IEA.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:006457; F:protein folding; IEA.
DR GO; GO:0006986; P:response to unfolded protein; IEA.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ_1.
DR PRINTS; PR00625; DNABPROTEIN.
DR SMART; SM00271; DnaJ_1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS50076; DnaJ_2; 1.
KM Chaperone; Heat shock.
SQ SEQUENCE 260 AA; 28639 MW; DDC645DAF9B41ED CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 260;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
```

Db 19 CAASAFSAC 28

## RESULT 23

Q57VE4\_ANOGA PRELIMINARY; PRT; 304 AA.

AC Q57VE4;  
 DT 01-FEB-2005 (TRENBLREL. 29, Created)  
 DT 01-FEB-2005 (TRENBLREL. 29, Last sequence update)  
 DE 01-FEB-2005 (TRENBLREL. 29, Last annotation update)  
 DE ENSANGP0000025853 (Fragment).  
 GN ORFNames=ENSANGG0000024781;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;  
 OC Anophelinae; Anopheles.  
 OC NCBI\_TaxId=180454;  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=PEST;  
 RC The Anopheles gambiae Sequence Committee;  
 RT "Anopheles gambiae re-annotation."  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=PEST;  
 RG The Anopheles gambiae Sequence Committee;  
 RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAB01008839; EAL41365.1; -; Genomic\_DNA.  
 CC EMBL; AAB01008839; EAL41365.1; -; Genomic\_DNA.  
 FT NON\_TER 1 304  
 SQ SEQUENCE 304 AA; 29530 MW; 1F55F0FB846B16B9 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 304;  
 Best Local Similarity 20.0%; Pred. No. 2.6;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
 Db 20 CSSTSTTVC 29

## RESULT 24

Q582B8\_9TRYP PRELIMINARY; PRT; 344 AA.

AC Q582B8;  
 DT 10-MAY-2005 (TRENBLREL. 30, Created)  
 DT 10-MAY-2005 (TRENBLREL. 30, Last sequence update)  
 DT 10-MAY-2005 (TRENBLREL. 30, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=TB927.5.2480;  
 OS Trypanosoma brucei.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 OC NCBI\_TaxId=5631;  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=GUTat10.1;  
 RA Shedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,  
 RA Shallom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,  
 RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,  
 RA Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,  
 RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;  
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=GUTat10.1;  
 RA El-Sayed N.M., Khalak H., Adams M.D.;  
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=GUTat10.1;  
 RA Haas B., Blandin G., El-Sayed N.;  
 RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC091655; AAX80451.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 344 AA; 37545 MW; A3925C9B2A48BD53 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 344;  
 Best Local Similarity 20.0%; Pred. No. 2.7;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
 Db 251 CSSSTASQSC 260

RESULT 25  
 Q57Z20\_9TRYP PRELIMINARY; PRT; 361 AA.

AC Q57Z20;  
 DT 10-MAY-2005 (TRENBLREL. 30, Created)  
 DT 10-MAY-2005 (TRENBLREL. 30, Last sequence update)  
 DT 10-MAY-2005 (TRENBLREL. 30, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=TB927.5.4290;  
 OS Trypanosoma brucei.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 OC NCBI\_TaxId=5631;  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=GUTat10.1;  
 RC Shedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,  
 RA Shallom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,  
 RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,  
 RA Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,  
 RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=GUTat10.1;  
 RA El-Sayed N.M., Khalak H., Adams M.D.;  
 RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=GUTat10.1;  
 RA Haas B., Blandin G., El-Sayed N.;  
 RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC116668; AAX80579.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 361 AA; 40083 MW; 95BC9BDCB3B9895 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 361;  
 Best Local Similarity 20.0%; Pred. No. 2.7;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13  
 Db 205 CSSSALSSSC 214

## RESULT 26

Q4YW87\_PLABE PRELIMINARY; PRT; 364 AA.

AC Q4YW87;  
 DT 13-SEP-2005 (TRENBLREL. 31, Created)  
 DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)  
 DT 13-SEP-2005 (TRENBLREL. 31, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 GN ORFNames=PB105184.00.0;  
 OS Plasmodium berghei.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OC NCBI\_TaxId=5821;

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Kairas M., Raine J.D., Carlton J.M., Koij T.W.A.,
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carnocci D.J., Yates J.R., Kafatos F.C.,
RA Jase C.J., Barrell B., Turner C.M.R., Waters A.P., Stinden R.S.,
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses."
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL: CAI01002130; CAH97719.1; -; Genomic_DNA.
CC Hypothetical protein.
CC NON TER
FT 1
SQ SEQUENCE 364 AA; 41304 MW; 1C1711BE940AA94F CRC64;

Query Match 53.1%; Score 17; DB 2; Length 364;
Best Local Similarity 20.0%; Pred. No. 2.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13
DB 332 CSTSTTAEBC 341

RESULT 27
Q70RD3_GERRY PRELIMINARY; PRT; 368 AA.
ID Q70RD3;
AC Q70RD3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MYB1 protein.
GN Name=Myb1;
OS Gerbera hybrid cv. 'Terra Regina'.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC asterids; campanulids; Asterales; Asteraceae; Mutistoidae; Mutistaceae;
OC Gerbera.
OC NCBI_Taxid=226891;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Petal;
RX PubMed=14605235; DOI=10.1104/pp.103.026039;
RA Elomaa P., Uimari A., Mehto M., Albert V.A., Laitinen R., Teeri T.H.;
RT "Activation of anthocyanin biosynthesis in Gerbera hybrida
RT (Asteraceae) suggests conserved protein-protein and protein-promoter
RT interactions between the anciently diverged monocots and eudicots."
RL Plant Physiol. 133:1831-1842(2003).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC EMBL: AJ554697; CAD87007.1; -; mRNA.
CC HSSP: 003237; 1A5J.
DR GO: 0005634; C:nucleus; IEA.
DR GO: 0005677; F:DNA binding; IEA.
DR InterPro: IPR012287; Homeodomain-rel.
DR InterPro: IPR01005; Myb DNA bd.
DR Pfam: PF00249; Myb DNA-binding; 2.
DR SMART: SM00717; SANT; 2.
DR PROSITE: PS00037; MYB_1; UNKOWN_1.
DR PROSITE: PS00334; MYB_2; 1.
DR PROSITE: PS50090; MYB_3; 2.
DR PROSITE: PS50090; MYB_3; 2.
KW Nuclear protein; Repeat.
SQ SEQUENCE 368 AA; 40911 MW; 7CFEFCBAD3AD9421 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 368;
Best Local Similarity 20.0%; Pred. No. 2.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

```

```

DB 286 CATSTTDSBC 295

RESULT 28
CBID_AGRTS
ID CBID_AGRTS STANDARD; PRT; 376 AA.
AC 08UBD6;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Putative cobalt-precortin-6A synthase (deacylating) (EC 2.1.1.-).
GN Name=cbid; OrderedLocustNames=Atu2795, AGP_C.5073.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Agrobacterium.
OC NCBI_Taxid=176299;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Secubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Bisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krepsan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chunley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard C.,
RA Grollio B., Goldman B.S., Cao Y., Askenzi M., Halling C., Mullin L.,
RA Hounel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughy D., Scott C., Jappes C., Markelz B.,
RA Flanagan C., Crowell C., Garson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
CC -!- FUNCTION: May catalyze the methylation of C-1 in cobalt-precortin-
CC 5 and the subsequent extrusion of acetic acid from the resulting
CC intermediate to form cobalt-precortin-6A.
CC -!- PATHWAY: Adenosylcobalamin biosynthesis; anaerobic branch of
CC corrin ring synthesis; seventh step.
CC -!- SIMILARITY: Belongs to the cbid family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AE009227; AAL43776.1; -; Genomic_DNA.
CC EMBL: AE008192; AAK88508.1; ALT_INIT; Genomic_DNA.
CC PIR: AB2920; AB2920.
CC PIR: C97694; C97694.
CC HAMAP: MF_00787; -; 1.
DR InterPro: IPR002748; Cbid.
DR Pfam: PF01888; Cbid; 1.
DR TIGRFAMs: TIGR00312; cbid; 1.
KW Cobalamin biosynthesis; Complete proteome; Transferase.
SQ SEQUENCE 376 AA; 39088 MW; D5437CB4CA463D67 CRC64;

Query Match 53.1%; Score 17; DB 1; Length 376;
Best Local Similarity 20.0%; Pred. No. 2.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13

```

Db 17 CAAATKAC 26

## RESULT 29

Q7PY13 ANOGA PRELIMINARY; PRT; 388 AA.

ID Q7PY13

DT 01-MAR-2004 (TrEMBLrel. 26, Created)

DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE ENSANGP0000018378 (Fragment).

GN ORFNames=ENSANG0000018589;

OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;

OC Anophelinae; Anopheles.

NCBI\_TaxId=180454;

NP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST.

RA Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

CC -! CAUTION: The sequence shown here is derived from an

EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

preliminary data.

DR EMBL; AAA01008987; EAA01136.1; -; Genomic\_DNA.

DR pfam; PF00566; TBC; 1. RabGAP\_TBC.

DR PROSITE; PS50086; TBC\_RABGAP; 1.

FT NON TER 388 388

SQ SEQUENCE 388 AA; 44534 MW; A3EEC567D458CCFF CRC64;

Query Match 53.1%; Score 17; DB 2; Length 388;  
Best Local Similarity 20.0%; Pred. No. 2.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 38 CSTSAFTC 47

## RESULT 30

Q21081 CAEEL PRELIMINARY; PRT; 389 AA.

ID Q21081

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Hypothetical protein K01C8.2.

GN ORFNames=K01C8.2;

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

NCBI\_TaxId=6239;

NP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=Bristol N2.

RX MEDLINE=99069613; PubMed=9851916;

RG The C. elegans sequencing consortium;

RT "Genome sequence of the nematode C. elegans: a platform for

investigating biology."

RL Science 282:2012-2018(1998).

DR EMBL; Z49068; CAA8855.1; -; Genomic\_DNA.

DR PIR; T23167; T23167.

DR Ensemble; K01C8.2; Caenorhabditis elegans.

DR Wormbase; WBGene00010457; K01C8.2.

DR WormPeP; K01C8.2; CE02263.

DR InterPro; IPR006150; Worm\_repeat\_1.

DR SMART; SM00289; WRI; 7.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 389 AA; 40335 MW; C32B3846991D17D CRC64;

Query Match 53.1%; Score 17; DB 2; Length 389;  
Best Local Similarity 20.0%; Pred. No. 2.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 123 CSSSISTSSC 132

## RESULT 31

Q67U23 ORYSA PRELIMINARY; PRT; 411 AA.

ID Q67U23

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Chloroplast thylakoidal processing peptidase-like protein.

GN Name=PO48D02.23;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

NCBI\_TaxId=39947;

NP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC

clone:PO48D02.23;"

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP005393; BAD38026.1; -; Genomic\_DNA.

DR Gramene; Q67U23; -.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR GO; GO:0008236; F:serine-type peptidase activity; IEA.

DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.

DR InterPro; IPR006198; Pept\_S24\_S26.

DR InterPro; IPR006198; Pept\_S24\_S26.

DR Pfam; PF00717; Peptidase\_S24; 1.

DR PRINTS; PR00727; LEADERPTASE.

DR TIGRPFAM; TIGR02227; sigrep\_1\_bact; 1.

DR PROSITE; PS00501; SPASE\_1\_1; 1.

DR PROSITE; PS00761; SPASE\_1\_3; 1.

KW Hydrolase.

SQ SEQUENCE 411 AA; 43573 MW; EE76AD564816A373 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 411;  
Best Local Similarity 20.0%; Pred. No. 2.9;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
DB 25 CSSASTAGC 34

## RESULT 32

Q41408 GIBZE PRELIMINARY; PRT; 437 AA.

ID Q41408

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=FG07800.1;

OS Gibberella zeae PH-1.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.

NCBI\_TaxId=229533;

NP NUCLEOTIDE SEQUENCE.

RC STRAIN=PH-1;

RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,

RA Arachchi H.M., Barta N., Bastien V., Bloom T., Boguslavsky L.,

RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,

RA Choepel Y., Collamore A., Cook A., Cooke P., Corum B., Dearellano K.,  
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,  
RA Erickson J., Fero S., Ferreira P., FitzGerald M., Gage D., Galagan J.,  
RA Garbaya S., Gierke S., Graham L., Grand-Pierre N., Hafiz N.,  
RA Hagopian D., Hago B., Hall J., Horton L., Hulme W., Iliev I.,  
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karates A.,  
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,  
RA Ma L.-U., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,  
RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,  
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,  
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,  
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,  
RA Rachtupia A., Ramsamy U., Raymond C., Retta R., Rise C., Rogov P.,  
RA Roman J., Schauer S., Schupack R., Seaman S., Severy P., Smirnov S.,  
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs W.,  
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,  
RA Vassiliev H., Venkataraman V.S., Vitel R., Vo A., Wang S., Wilson B.,  
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,  
RA Lander E.;  
RA "Fusarium graminearum genome sequence."  
RA Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AACN01000320; EAA77994.1; -; Genomic DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 437 AA; 47817 MW; 01EBA0260503394 CRC64;  
  
Query Match 53.1%; Score 17; DB 2; Length 437;  
Best Local Similarity 20.0%; Pred. No. 2.9;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
OY 4 CXXXXXXXXX 13  
DB 199 CATLASTSSC 208  
  
RESULT 33  
Q9VMG7\_DROME PRELIMINARY; PRT; 471 AA.  
ID Q9VMG7\_DROME PRELIMINARY; PRT; 471 AA.  
AC Q9VMG7;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE CG13990-PA.  
OS Name=CG13990; ORFNames=CG13990;  
GN Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amarantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blaziel R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.F., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borikova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Donip L.E., Domene M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster A., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.B., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,  
RA Liao P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Murray D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacleb J.M.,  
RA Palczolo M., Pittman G.S., Pan S., Pollard J., Ruit V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs S.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RT Science 287:2185-2195(2000).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celisner S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
RT melanogaster euchromatic genome sequence."  
RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,  
RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celisner S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
RT a genome perspective."  
RT Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Mitra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield B.J., Bayraktaroglu L., Berman B.P.,  
RA Betencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review."  
RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RX Berkeley Drosophila Genome Project;  
RA Celisner S., Carlson J., Wan K., Pfeiffer B., Friese E., George R.,  
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,  
RA Yu C., Rubin G.;  
RT "Drosophila melanogaster release 4 sequence."  
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RX Flybase;  
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AEO03612; AAF52351.2; -; Genomic DNA.  
RL Ensemble; CG13990; Drosophila melanogaster.  
DR Flybase; FBgn0040950; CG13990.  
DR GO; GO:0005576; C:extracellular region; IEA.  
DR GO; GO:0008061; P:chitin binding; IEA.  
DR GO; GO:0006030; P:chitin metabolism; IEA.  
DR InterPro; IPR002557; Chitin\_bind\_Pera.  
DR Pfam; PF01607; CBM\_14; 2.

DR SMART: SM00494; Chcbp2; 2.  
 DR PROSITE: PSS0940; CHIT BIND II; 2.  
 SQ SEQUENCE 471 AA; 4811 MW; F3167EB0728E9578 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 471;  
 Best Local Similarity 20.0%; Pred. No. 3;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 DB 350 CASSTTTEC 359

RESULT 34  
 Q7X7A4\_ORYSA PRELIMINARY; PRT; 475 AA.

AC Q7X7A4;  
 DT 01-OCT-2003 (TrEMBLrel. 25, last sequence update)  
 DT 10-OCT-2003 (TrEMBLrel. 30, last annotation update)  
 DE OSJNB0060P14.14 protein (OSJNB0048E02.5 protein).  
 GN Name=OSJNB0060P14.14; SYNONYMS=OSJNB0048E02.5;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 NC NCB1\_TaxID=39947;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RX PubMed=12447439; DOI=10.1038/nature01183;  
 RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,  
 RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Yang K., Yu S., Tang J.,  
 RA Weng Q., Zhang L., Lu Y., Wu J., Lu Y., Zhang L.S., Yu Z., Fan D.,  
 RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,  
 RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,  
 RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,  
 RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,  
 RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,  
 RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,  
 RA Han B.;  
 RT "Sequence and analysis of rice chromosome 4.";  
 RL Nature 420:316-320(2002).  
 DR EMBL: AL663017; CA604361.2; -; Genomic\_DNA.  
 DR EMBL: AL606653; CA604825.2; -; Genomic\_DNA.  
 DR Gramene: Q7X7A4; -;  
 DR GO: GO:006512; P:ubiquitin cycle; IEA.  
 DR InterPro: IPR001810; F-box.  
 DR Pfam: PF00646; F-box; 1.  
 DR SMART: SM00256; FBOX; 1.  
 DR PROSITE: PSS0181; FBOX; 1.  
 SQ SEQUENCE 475 AA; 53063 MW; 63345B3CB2D4CA5 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 475;  
 Best Local Similarity 20.0%; Pred. No. 3;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 DB 417 CSSTLSAC 426

RESULT 35  
 Q4P4C7\_USTMA PRELIMINARY; PRT; 491 AA.

AC Q4P4C7;  
 DT 13-SEP-2005 (TrEMBLrel. 31, last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, last sequence update)  
 DE Hypothetical protein.  
 DE ORFNames=UW05036.1;  
 GN Ustilago maydis 521.  
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;  
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.

OX NCB1\_TaxID=237631;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=521.  
 RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adkoya E.,  
 RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,  
 RA Arachchi H., Armbruster J., Bachantansy P., Baldwin J., Barry A.,  
 RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L.,  
 RA Borowsky M., Bouhgalter B., Brumache A., Butler J., Calixte N.,  
 RA Calvo S., Camarata J., Campo K., Chang J., Cheebatang Y., Chtren M.,  
 RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,  
 RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,  
 RA Dorjee K., Dorris L., Duffey N., Dupes A., Ekins T., Engels R.,  
 RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,  
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Gartin G., Gnerre S.,  
 RA Gnirke A., Goette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,  
 RA Hagopian D., Hages B., Hall J., Hatcher B., Heller A., Higgins H.,  
 RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,  
 RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysseilis M., Karlsson E.,  
 RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutti K.,  
 RA Lama D., Landers T., Leger J., Levins S., Lewis D., Lewis T.,  
 RA Lindblad-toh K., Liu X., Lokysang T., Lokysang Y., Lucien O.,  
 RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,  
 RA Manning J., Mardella R., Maru K., Matthews C., Mauceli E.,  
 RA McCarthy M., McDonough S., Mcghee T., Meldrum J., Meneus L.,  
 RA Mestrov J., Mihalev A., Mihova T., Mikkelson T., Mlenga V., Moru K.,  
 RA Mozes J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,  
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,  
 RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,  
 RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignat B.,  
 RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,  
 RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,  
 RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe C.,  
 RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnuez C.,  
 RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,  
 RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchinga P.,  
 RA Tenzing P., Tesfaye S., Theodore J., Thoulutang Y., Topham K.,  
 RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliou H.,  
 RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,  
 RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,  
 RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembek L.,  
 RA Zimmer A., Zody M., Zander B.;  
 RT "The genome sequence of Ustilago maydis.";  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL: AACP0100186; EAK85896.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 491 AA; 51869 MW; 760A84FBB8D71BBE CRC64;

Query Match 53.1%; Score 17; DB 2; Length 491;  
 Best Local Similarity 20.0%; Pred. No. 3.1;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXX 13  
 DB 453 CSASTGSSC 462

RESULT 36  
 Q4XZV5\_PLACH PRELIMINARY; PRT; 520 AA.

AC Q4XZV5;  
 DT 13-SEP-2005 (TrEMBLrel. 31, last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, last sequence update)  
 DE Hypothetical protein (Fragment).  
 DE ORFNames=PC000338.02.0;  
 GN Plasmodium chabaudi.  
 OS Plasmodium chabaudi.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCB1\_TaxID=5825;  
 RN (1)



RP NUCLEOTIDE SEQUENCE.  
RA Hall N., Karris M., Raine J.D., Carlton J.M., Kooij T.W.A.,  
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,  
RA James K., Rutherford K., Harris B., Harris D., Church C.,  
RA Quail M., Ormond D., Doggett J., Trueman H.E., Mendoza J.,  
RA Bidwell S.L., Rajandream M.A., Carrucci D.J., Yates J.R., Kafatos F.C.,  
RA Jane C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.,  
RT "A comprehensive survey of the Plasmodium life cycle by genomic,  
RT transcriptomic, and proteomic analyses."  
RL Science 307:82-86(2005).  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL: CAJ01002036; CANT7555.1; -: Genomic\_DNA.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00400; WD40; 5.  
DR PRINTS: PR00320; GPROTEINBRPT.  
DR SMART: SM00320; WD40; 5.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_1.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Hypothetical protein; Repeat; WD repeat.  
KW NON\_TER 520 520  
SQ SEQUENCE 520 AA; 59951 MW; FA789C11BFD99087 CRC64;  
  
Query Match 53.1%; Score 17; DB 2; Length 520;  
Best Local Similarity 20.0%; Pred. No. 3.1;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 4 CXXXXXXXXX 13  
DB 364 CSTSSSSSC 373  
  
RESULT 37  
ID 0803D5 BRARE PRELIMINARY; PRT; 556 AA.  
AC 0803D5  
DT 01-UN-2003 (TREMBLrel. 24, Created)  
DT 01-UN-2003 (TREMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE Similar to adrenal secretory serine protease.,  
GN ORFNames=zgc:55888;  
OS Brachydanio rerio (zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OC NCBI\_Taxid=7955;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN-AB; TISSUE=whole body;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stabileton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein W.J., Ushed T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.C., Touchman J.W., Green B.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marrs M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.

RC STRAIN-AB; TISSUE=whole body;  
RA Strausberg R.,  
RT Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC044526; AAH44526.1; -: mRNA.  
DR HSSP: P00760; IEZK.  
DR Ensembl: ENSDARG00000016538; Danio rerio.  
DR ZFIN: ZDB-GENE-040426-834; zgc:55888.  
DR GO: GO:0004263; F-tyrosinase activity; IEA.  
DR GO: GO:0008233; F-peptidase activity; IEA.  
DR GO: GO:0044295; F-tyrosinase activity; IEA.  
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro: IPR001254; Peptidase\_S1\_S6.  
DR InterPro: IPR001314; Peptidase\_S1A.  
DR Pfam: PF00089; Trypsin\_2.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR SMART: SM00020; Tryp\_Spc; 2.  
DR PROSITE: PS50240; TRYPsin\_DOM; 2.  
DR PROSITE: PS00134; TRYPsin\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00135; TRYPsin\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
KW SEQUENCE 556 AA; 61077 MW; C7B23D930547878C CRC64;  
  
Query Match 53.1%; Score 17; DB 2; Length 556;  
Best Local Similarity 20.0%; Pred. No. 3.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 4 CXXXXXXXXX 13  
DB 493 CTHAAASTSC 502  
  
RESULT 38  
ID 05GXM1\_XANOR PRELIMINARY; PRT; 568 AA.  
AC 05GXM1;  
DT 10-MAY-2005 (TREMBLrel. 30, Created)  
DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)  
DE Alkaline phosphatase.  
GN Name=phoA; OrderedLocustNames=XO03296;  
OS Xanthomonas oryzae (pv. oryzae).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xanthomonas.  
OC NCBI\_Taxid=64187;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=XACC10331 / KX085;  
RX PubMed=15673718; DOI=10.1093/nar/gki206;  
RA Lee B.-M., Park Y.-U., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,  
RA Park I.-C., Yoon U.-H., Hahn J.-H., Ko B.-S., Lee G.-B., Kim H.,  
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo U.-S.,  
RA Go S.-U.;  
RT "The genome sequence of Xanthomonas oryzae pathovar XACC10331,  
RT the bacterial blight pathogen of rice."  
RL Nucleic Acids Res. 33:577-586(2005).  
DR EMBL: AB013598; AAW76550.1; -: Genomic\_DNA.  
DR GO: GO:0008152; P:metabolism; IEA.  
DR InterPro: IPR001952; ALK\_phosphatase.  
DR Pfam: PF00245; ALK\_phosphatase; 1.  
DR PRINTS: PR00113; ALKPHPTASE.  
DR SMART: SM00098; alkPpc; 1.  
KW Complete proteome.  
SQ SEQUENCE 568 AA; 60106 MW; CAE69E266ACB3A6 CRC64;  
  
Query Match 53.1%; Score 17; DB 2; Length 568;  
Best Local Similarity 20.0%; Pred. No. 3.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 4 CXXXXXXXXX 13  
DB 21 CASTAGSSAC 30



```

RESULT 39
OAN2Y7 THEPA PRELIMINARY; PRT; 626 AA.
ID OAN2Y7_
AC OAN2Y7_
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Hypothetical protein.
OC ORFNames=TP04_0205;
OS Theileria parva.
OC Eukaryota, Alveolata, Apicomplexa, Piroplasmida, Theileridae;
OC Theileria.
OX NCBI_TaxID=5875;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Muga;
RA Gardner M.D., Bishop R., Shah T., de Villiers E.P., Carlton J.M.,
RA Hall N., Ren Q., Paulsen I.T., Pain A., Berriman M., Wilson R.J.M.,
RA Sato S., Ralph S.A., Mann D.J., Xiong Z., Shallom S.J., Weidman J.,
RA Jiang L., Lynn J., Weaver B., Shoabdi A., Maawo D., Crabtree J.,
RA Mortman J.R., Haas B., Angiolini S.V., Creasy T.H., Lu C., Suh B.,
RA Silva J.C., Uterback T.R., Feldblyum T.V., Perrea M., Allen J.,
RA Taracha E.L.N., Salzberg S.L., White O.R., Fitzhugh H.A., Morzaria S.,
RA Venter J.C., Fraser C.M., Nene V.;
RT "Genome sequence of Theileria parva, a bovine pathogen that transforms
RT lymphocytes."
RT Science 309:134-137(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Muga;
RA Gardner M., Bishop R., Shah T., de Villiers E., Carlton J.M., Hall N.,
RA Ren Q., Paulsen I.T., Pain A., Berriman M., Wilson R.J.M., Sato S.,
RA Ralph S.A., Mann D.J., Xiong Z., Shallom S.J., Weidman J., Jiang L.,
RA Lynn J., Weaver B., Shoabdi A., Maawo D., Crabtree J., Mortman J.R.,
RA Haas B., Angiolini S., Creasy T.H., Lu C., Suh B., Silva J.C.,
RA Uterback T., Feldblyum T., Perrea M., Allen J., Taracha E.L.,
RA Salzberg S.L., White O., Fitzhugh H.A., Morzaria S., Venter J.C.,
RA Fraser C.M., Nene V.;
RT Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAGK01000004; EMBL557.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 626 AA; 72615 MW; 663085BA9509916 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 626;
Best Local Similarity 20.0%; Pred. No. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 13
DB 44 CAASSSDASC 53

RESULT 40
O69TY8 ORYSA PRELIMINARY; PRT; 630 AA.
ID O69TY8_
AC O69TY8_
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
DE Receptor protein kinase-like.
GN Name=OSUNBA0029G06.34; Synonym=OSUNBA0072A21.3;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC

```

```

RT clone:OSUNBA0029G06.";
RN Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
RT clone:OSUNBA0072A21.";
RN Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004680; BAD35689.1; -; Genomic_DNA.
DR EMBL; AP004737; BAD37734.1; -; Genomic_DNA.
DR Gramene; O69TY8; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR GO; GO:0004668; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR002482; LysM.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF01476; LysM; 1.
DR Pfam; PF00069; PKinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
KW kinase, Receptor.
SQ SEQUENCE 630 AA; 67017 MW; F0EFC8DB8A63646 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 630;
Best Local Similarity 20.0%; Pred. No. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXC 13
DB 48 CSEYTTATGTC 57

RESULT 41
HSP71 LEIMA STANDARD; PRT; 634 AA.
ID HSP71_LEIMA
AC P12076;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial precursor.
GN Name=HSP70.1;
OS Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MHOM/IL/81/Friedlin;
RX MEDLINE=93300981; PubMed=8314893;
RA Seattle S., McCrosan M.V., Smith D.F.;
RT "Expression of a mitochondrial stress protein in the protozoan
RT parasite Leishmania major.";
RL J. Cell Sci. 104:1091-1100(1993).
RN [2]
RP SEQUENCE REVISION TO 461-500.
RA Smith D.F.;
RN Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-249.
RC STRAIN=MHOM/IL/81/Friedlin;
RX MEDLINE=89345072; PubMed=2762121;
RA Seattle S., Campos A.J.R., Coulson R.M.R., Spithill T.W., Smith D.F.;
RT "A family of heat shock protein 70-related genes are expressed in the
RT promastigotes of Leishmania major.";
RL Nucleic Acids Res. 17:5081-5095(1989).
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.

```

```
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL, X64137; CAA5498.2; -; Genomic_DNA.
DR EMBL, X14574; CAA32713.1; -; Genomic_DNA.
DR PIR, S33575; S33575.
DR HSSP, P04475; 1DKG.
DR InterPro, IPR012725; DnaK prot.
DR InterPro, IPR001023; Hsp70.
DR PANTHER, PTHR19375; Hsp70; 1.
DR Pfam, PF00012; HSP70; 1.
DR PRINTS, PR00301; HEATSHOCK70.
DR ProDom, PD000089; Hsp70; 1.
DR TIGRFAMs, TIGR02350; prok_dnaK; 1.
DR PROSITE, PS00297; HSP70_1; 1.
DR PROSITE, PS00329; HSP70_2; 1.
DR PROSITE, PS01036; HSP70_3; FALSE_NEG.
DR ATP-binding; Coiled coil; Heat shock; Mitochondrion; Multigene family;
DR Nucleotide-binding; Transic peptide.
FT TRANSIT 1 634 Mitochondrion (By similarity).
FT CHAIN 20 Heat shock 70-related protein 1.
FT COILED 538 614 Potential.
SQ SEQUENCE 634 AA; 68330 MW; 8382B1C40C108BEA CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 1; Length 634;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 7 CGSAAASAAAC 16

RESULT 42
Q04747_LEIMA PRELIMINARY; PRT; 635 AA.
ID Q04747;
AC Q04747;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=lmj30.2460;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Fredlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL, CT005267; CAJ06531.1; -; Genomic_DNA.
DR InterPro, IPR012725; DnaK prot.
DR InterPro, IPR001023; Hsp70.
DR Pfam, PF00012; HSP70; 1.
DR PRINTS, PR00301; HEATSHOCK70.
DR ProDom, PD000089; Hsp70; 1.
DR TIGRFAMs, TIGR02350; prok_dnaK; 1.
DR PROSITE, PS00297; HSP70_1; 1.
DR PROSITE, PS00329; HSP70_2; 1.
DR PROSITE, PS01036; HSP70_3; 1.
DR ATP-binding; Heat shock; Nucleotide-binding.
SQ SEQUENCE 635 AA; 68948 MW; 1A11F63117CA2200 CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 635;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 4 CXXXXXXXXXC 13
DB 7 CGSAAASAAAC 16

RESULT 43
Q0740_LEIMA PRELIMINARY; PRT; 652 AA.
ID Q0740;
AC Q0740;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=lmj30.2550;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Fredlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL, CT005267; CAJ06551.1; -; Genomic_DNA.
DR InterPro, IPR012725; DnaK prot.
DR InterPro, IPR001023; Hsp70.
DR Pfam, PF00012; HSP70; 1.
DR PRINTS, PR00301; HEATSHOCK70.
DR ProDom, PD000089; Hsp70; 1.
DR TIGRFAMs, TIGR02350; prok_dnaK; 1.
DR PROSITE, PS00297; HSP70_1; 1.
DR PROSITE, PS00329; HSP70_2; 1.
DR PROSITE, PS01036; HSP70_3; 1.
DR ATP-binding; Heat shock; Nucleotide-binding.
SQ SEQUENCE 652 AA; 70609 MW; 52D4843JDEC36001 CRC64;

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 652;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 7 CGSAAASAAAC 16

RESULT 44
Q0744_LEIMA PRELIMINARY; PRT; 660 AA.
ID Q0744;
AC Q0744;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=lmj30.2490;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Fredlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL, CT005267; CAJ06539.1; -; Genomic_DNA.
DR InterPro, IPR012725; DnaK prot.
DR InterPro, IPR001023; Hsp70.
DR Pfam, PF00012; HSP70; 1.
DR PRINTS, PR00301; HEATSHOCK70.
```

DR ProDom; PD000089; Hsp70; 1.  
 DR TIGRAME; TIGR02350; prok\_dnak; 1.  
 DR PROSITE; PS00297; HSP70\_1; 1.  
 DR PROSITE; PS00329; HSP70\_2; 1.  
 DR PROSITE; PS01036; HSP70\_3; 1.  
 DR ATP-binding; Heat shock; Nucleotide-binding.  
 KW ATP-binding; Heat shock; Nucleotide-binding.  
 SQ SEQUENCE 660 AA; 71621 MW; 5E1D306B3D524242 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 660;  
 Best Local Similarity 20.0%; Pred. No. 3.4;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXC 13  
 Db 7 CGSAASAAAC 16

RESULT 45  
 ID Q40745 LEIMA PRELIMINARY; PRT; 662 AA.

AC Q40745-1  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Heat shock 70-related protein 1, mitochondrial, putative.  
 GN ORFAMES-LmjF30.2470, LmjF30.2480;  
 OS Leishmania major.

OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_TaxID=5664;  
 RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Friedlin;

RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,  
 RA Smith D., Collins M., Rosker N., Harris D., Oliver K., O'Neill S.,  
 RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;  
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.

CC -1 SIMILARITY: Belongs to the heat shock protein 70 family.

DR EMBL; CT005267; CAJ06535.1; -; Genomic DNA.

DR EMBL; CT005267; CAJ06535.1; -; Genomic DNA.

DR InterPro; IPR012725; Dnak\_prok.

DR InterPro; IPR001023; Hsp70.

DR Pfam; PF00012; HSP70; 1.

DR PRINTS; PR00301; HEATSHOCK70.

DR ProDom; PD000089; Hsp70; 1.

DR TIGRAME; TIGR02350; prok\_dnak; 1.

DR PROSITE; PS00297; HSP70\_1; 1.

DR PROSITE; PS00329; HSP70\_2; 1.

DR PROSITE; PS01036; HSP70\_3; 1.

KW ATP-binding; Heat shock; Nucleotide-binding.

SQ SEQUENCE 662 AA; 71877 MW; CFC66508C46E20 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 662;  
 Best Local Similarity 20.0%; Pred. No. 3.4;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXC 13  
 Db 7 CGSAASAAAC 16

RESULT 46  
 ID Q60E18 ORYSA PRELIMINARY; PRT; 691 AA.

AC Q60E18-1

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Putative cyclic nucleotide gated ion channel.

GN Name=OSUNBA0017K09.10;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI\_TaxID=39947;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,

RA Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,

RA Chao Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,

RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kuo P.-I., Lee M.-C.,

RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,

RA Wu H.-P., Shaw J.-F.,

RT "Oryza sativa BAC OSUNBA0017K09 genomic sequence."

RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC130597; AAU90233.1; -; Genomic DNA.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005216; F:ion channel activity; IEA.

DR GO; GO:0006811; P:ion transport; IEA.

DR InterPro; IPR000595; cNMP\_binding.

DR InterPro; IPR005821; Ion\_trans.

DR Pfam; PF00027; cNMP\_binding; 1.

DR Pfam; PF00520; Ion\_trans; 1.

DR SMART; SM00100; cNMP; 1.

DR PROSITE; PS00042; cNMP\_BINDING\_3; 1.

SQ SEQUENCE 691 AA; 77220 MW; D059333316F2E299 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 691;  
 Best Local Similarity 20.0%; Pred. No. 3.5;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXC 13  
 Db 298 CASSTAFAVAC 307

RESULT 47  
 ID Q819P8 9BIVA PRELIMINARY; PRT; 699 AA.

AC Q819P8-1

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Alpha-amylase.

GN Name=Amv;

OS Corbicula fluminea.

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;

OC Corbiculoidea; Corbiculidae; Corbicula.

OX NCBI\_TaxID=45949;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Da Lage J.-L., Van Wormhoudt A., Cariou M.-L.,

RT "Diversity and evolution of the alpha-amylase genes in Animals."

RL Biologia 57:181-189(2002).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RA PubMed=14704857; DOI=10.1007/s00018-003-3334-Y;

RT Da Lage J.-L., Feller G., Janacek S.,

RT "Horizontal gene transfer from Eukarya to bacteria and domain

RT shuffling: the alpha-amylase model."

RL Cell. Mol. Life Sci. 61:97-109(2004).

RN [3]

RP NUCLEOTIDE SEQUENCE.

RA Da Lage J.-L.,

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF468016; AA017927.2; -; Genomic DNA.

DR HSSP; P04745; LSMD.

DR GO; GO:0004556; F:alpha-amylase activity; IEA.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR GO; GO:0016798; F:hydrolase activity; acting on glycosyl bonds; IEA.

DR GO; GO:0005975; P:carbohydrate metabolism; IEA.

DR InterPro; IPR006048; Alpha\_ami1\_C.

DR InterPro; IPR006048; Alpha\_ami1\_C.

DR InterPro; IPR006048; Glyco\_hydro\_13.

DR Pfam; PF00128; Alpha-amylase; 1.

DR Pfam; PF02806; Alpha-amylase\_C; 1.

```

DR PRINTS: PR00110; ALPHAMYLASE.
DR SMART: SM00642; Amy; 1.
DR SMART: SM00632; Amy; C; 1.
SQ SEQUENCE 699 AA; 76544 MW; 25D57008165B04CB CRC64;

Query Match 53.1%; Score 17; DB 2; Length 699;
Best Local Similarity 20.0%; Pred. No. 3.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 544 CTSDATSAC 553

RESULT 48
O5L6X2 CHLAB
ID O5L6X2 CHLAB PRELIMINARY; PRT; 705 AA.
AC O5L6X2-
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Putative transport protein.
GN OrderedLocNames=CAB140;
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OK NCBI_TaxId=83555;
RN NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RP STRAIN=S26/3;
RC PubMed=15837807; DOI=10.1101/gr.3684805;
RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,
RA Livingstone M., Cerdano-Tarraga A.-M., Harris B., Doggett J.,
RA Ormond D., Mungall K., Clarke K., Fellwell T., Hance Z., Sanders M.,
RA Quail M.A., Price C., Barrett B.G., Parkhill J., Longbottom D.;
RT "The Chlamydomophila abortus genome sequence reveals an array of
RT variable proteins that contribute to interspecies variation.";
RL Genome Res. 15:629-640(2005).
DR EMBL; CR846038; CAH63598.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 705 AA; 81369 MW; 1997374B5266E2AB CRC64;

Query Match 53.1%; Score 17; DB 2; Length 705;
Best Local Similarity 20.0%; Pred. No. 3.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 67 CSSLSATTC 76

RESULT 49
O9NKE3 DROME
ID O9NKE3 DROME PRELIMINARY; PRT; 734 AA.
AC O9NKE3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Hypothetical protein BG:DS00180.5.
GN Name=BG:DS00180.5; ORFNames=CG31766, CG31766, CG31841;
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OK NCBI_TaxId=7227;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Berkley;
RC MEDLINE=99403001; PubMed=10471707;
RA Ashburner M., Mitra S., Roote J., Lewis S.E., Blazej R.G., Davis T.,
RA Doyle C., Galle R.F., George R.A., Harris N.L., Hartzell G.,
RA Harvey D.A., Hong L., Houston K.A., Hoskins R.A., Johnson G.,
RA Martin C., Moshrefi A.R., Palazzo M., Reese M.G., Spradling A.C.,
RA Tsang G., Wan K.H., Whitelaw K., Celniker S.E., Rubin G.M.;

```

```

RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummel S.R., Katta K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomolan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Smit E., Swirek R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003408; AAF44838.1; -; Genomic_DNA.
DR FlyBase; FBgn0051765; CG31765.
DR FlyBase; FBgn0051766; CG31766.
DR FlyBase; FBgn0051841; CG31841.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004246; F:peptidyl-dipeptidase A activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001548; Peptidase_M2.
DR Pfam; PF01401; Peptidase_M2; 1.
KW Hypothetical protein.
SQ SEQUENCE 734 AA; 79831 MW; 4C0B5993AF9FBC1E CRC64;

Query Match 53.1%; Score 17; DB 2; Length 734;
Best Local Similarity 20.0%; Pred. No. 3.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 681 CSSASAVASC 690

RESULT 50
Q4YNR2_PLABE
ID Q4YNR2_PLABE PRELIMINARY; PRT; 784 AA.
AC Q4YNR2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=PB000891.03.0;
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OK NCBI_TaxId=5821;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karris M., Raine J.D., Carlton J.M., Kool J.T.W.A.,
RA Bertman M., Florens L., Jansen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Jance C.J., Barrett B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAI01003291; CAI00348.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 784 AA; 91429 MW; B76DABAD7A0C1D6 CRC64;

Query Match 53.1%; Score 17; DB 2; Length 784;
Best Local Similarity 20.0%; Pred. No. 3.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 13
DB 363 CSSLSSSSC 372

```

Thu Jan 5 09:14:21 2006

us-09-932-322-4.rup

Page 25

Search completed: January 4, 2006, 16:09:24  
Job time : 131.487 secs

2006-01-04 16:09:24

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 ; Search time 123.183 Seconds  
(without alignments)  
64.204 Million cell updates/sec

Title: US-09-932-322-5  
Perfect score: 34  
Sequence: 1 XXXXXXXXXXXXXXXX 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : A\_Geneseq\_21:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*
- 9: Geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	52.9	104	7	Ab082562 Pseudomon
2	18	52.9	110	8	Adp30723 Human sec
3	18	52.9	135	8	Adp31638 Human sec
4	18	52.9	142	7	Ab079745 Pseudomon
5	18	52.9	178	7	Ab073284 Pseudomon
6	18	52.9	264	8	Adp31527 Human sec
7	18	52.9	270	8	Adp31435 Human sec
8	18	52.9	307	4	AbB65879 Drosophil
9	18	52.9	357	8	Adp31123 Human sec
10	18	52.9	615	8	Adp31132 Human sec
11	18	52.9	627	6	Abu45809 Pseudomon
12	18	52.9	627	4	Abu45809 Pseudomon
13	18	52.9	669	8	Adp31142 Human sec
14	18	52.9	711	8	Adp31125 Human sec
15	18	52.9	739	8	Adp31196 Human sec
16	18	52.9	771	8	Adp31244 Human sec
17	18	52.9	876	8	Adp31220 Human sec
18	18	52.9	882	8	Adp31688 Human sec
19	18	52.9	925	5	Adp31426 Human sec
20	18	52.9	1064	8	Adp31517 Human sec
21	18	52.9	1086	8	Adp31175 Human sec
22	18	52.9	1113	8	Adp31508 Human sec
23	18	52.9	1128	6	Adp315725 C. eleg
24	18	52.9	1134	8	Adp30741 Human sec

25	18	52.9	1134	8	Adp30924 Human sec
26	18	52.9	1168	8	Adp31046 Human sec
27	18	52.9	1191	8	Adp30993 Human sec
28	18	52.9	1260	8	Adp31533 Human sec
29	18	52.9	1289	8	Adp30675 Human sec
30	18	52.9	1437	8	Adp31357 Human sec
31	18	52.9	1454	8	Adp31177 Human sec
32	18	52.9	1480	8	Adp30557 Human sec
33	18	52.9	1588	5	AbB09437 H. influe
34	18	52.9	1617	8	Adp30660 Human sec
35	18	52.9	1652	6	Adp315715 C. eleg
36	18	52.9	1725	8	Adp30654 Human sec
37	18	52.9	1833	8	Adp30642 Human sec
38	18	52.9	2088	8	Adp31178 Human sec
39	18	52.9	2127	8	Adp31327 Human sec
40	18	52.9	2484	8	Adp30659 Human sec
41	18	52.9	2508	8	Adp315721 Human sec
42	18	52.9	2544	6	Adp315717 C. eleg
43	18	52.9	2601	6	Adp31299 Human sec
44	18	52.9	2833	8	Adp30572 Human sec
45	18	52.9	3411	8	Adp30657 Human sec
46	18	52.9	3411	8	Adp30657 Human sec
47	18	52.9	3411	8	Adp30657 Human sec
48	18	52.9	3411	8	Adp30657 Human sec
49	18	52.9	3411	8	Adp30657 Human sec
50	18	52.9	3411	8	Adp30657 Human sec
51	18	52.9	3411	8	Adp30657 Human sec
52	18	52.9	3411	8	Adp30657 Human sec
53	18	52.9	3411	8	Adp30657 Human sec
54	18	52.9	3411	8	Adp30657 Human sec
55	18	52.9	3411	8	Adp30657 Human sec
56	18	52.9	3411	8	Adp30657 Human sec
57	18	52.9	3411	8	Adp30657 Human sec
58	18	52.9	3411	8	Adp30657 Human sec
59	18	52.9	3411	8	Adp30657 Human sec
60	18	52.9	3411	8	Adp30657 Human sec
61	18	52.9	3411	8	Adp30657 Human sec
62	18	52.9	3411	8	Adp30657 Human sec
63	18	52.9	3411	8	Adp30657 Human sec
64	18	52.9	3411	8	Adp30657 Human sec
65	18	52.9	3411	8	Adp30657 Human sec
66	18	52.9	3411	8	Adp30657 Human sec
67	18	52.9	3411	8	Adp30657 Human sec
68	18	52.9	3411	8	Adp30657 Human sec
69	18	52.9	3411	8	Adp30657 Human sec
70	18	52.9	3411	8	Adp30657 Human sec
71	18	52.9	3411	8	Adp30657 Human sec
72	18	52.9	3411	8	Adp30657 Human sec
73	18	52.9	3411	8	Adp30657 Human sec
74	18	52.9	3411	8	Adp30657 Human sec
75	18	52.9	3411	8	Adp30657 Human sec
76	18	52.9	3411	8	Adp30657 Human sec
77	18	52.9	3411	8	Adp30657 Human sec
78	18	52.9	3411	8	Adp30657 Human sec
79	18	52.9	3411	8	Adp30657 Human sec
80	18	52.9	3411	8	Adp30657 Human sec
81	18	52.9	3411	8	Adp30657 Human sec
82	18	52.9	3411	8	Adp30657 Human sec
83	18	52.9	3411	8	Adp30657 Human sec
84	18	52.9	3411	8	Adp30657 Human sec
85	18	52.9	3411	8	Adp30657 Human sec
86	18	52.9	3411	8	Adp30657 Human sec
87	18	52.9	3411	8	Adp30657 Human sec
88	18	52.9	3411	8	Adp30657 Human sec
89	18	52.9	3411	8	Adp30657 Human sec
90	18	52.9	3411	8	Adp30657 Human sec
91	18	52.9	3411	8	Adp30657 Human sec
92	18	52.9	3411	8	Adp30657 Human sec
93	18	52.9	3411	8	Adp30657 Human sec
94	18	52.9	3411	8	Adp30657 Human sec
95	18	52.9	3411	8	Adp30657 Human sec
96	18	52.9	3411	8	Adp30657 Human sec
97	18	52.9	3411	8	Adp30657 Human sec

98	17	50.0	325	7	ABM87858	Abm87858	Rice	abio	171	17	50.0	950	8	ADP31167	Adp31167	Human	sec
99	17	50.0	330	7	ADB70062	Adb70062	C. neotor		172	17	50.0	951	8	ADP30943	Adp30943	Human	sec
100	17	50.0	348	8	ADP31441	Adp31441	Human	sec	173	17	50.0	957	8	ADP31528	Adp31528	Human	sec
101	17	50.0	357	8	ADP31267	Adp31267	Human	sec	174	17	50.0	957	8	ADP07508	Adp07508	Human	col
102	17	50.0	367	8	ADP30950	Adp30950	Human	sec	175	17	50.0	960	8	ADP31471	Adp31471	Human	sec
103	17	50.0	381	8	ADP30655	Adp30655	Human	sec	176	17	50.0	960	8	ADP31470	Adp31470	Human	sec
104	17	50.0	384	8	ADP30656	Adp30656	Human	sec	177	17	50.0	981	8	ADP30547	Adp30547	Human	sec
105	17	50.0	392	8	ADP31061	Adp31061	Human	sec	178	17	50.0	996	8	ADP31538	Adp31538	Human	sec
106	17	50.0	392	8	ADP31055	Adp31055	Human	sec	179	17	50.0	1035	8	ADP31552	Adp31552	Human	sec
107	17	50.0	393	8	ADP31345	Adp31345	Human	sec	180	17	50.0	1038	8	ADP30860	Adp30860	Human	sec
108	17	50.0	402	8	ADP30742	Adp30742	Human	sec	181	17	50.0	1044	8	ADP31550	Adp31550	Human	sec
109	17	50.0	402	8	ADP31414	Adp31414	Human	sec	182	17	50.0	1048	8	ADP31642	Adp31642	Human	sec
110	17	50.0	402	8	ADP31262	Adp31262	Human	sec	183	17	50.0	1053	8	ADP30886	Adp30886	Human	sec
111	17	50.0	409	7	ABO79810	AbO79810	Pseudomon		184	17	50.0	1059	8	ADP31042	Adp31042	Human	sec
112	17	50.0	411	8	ADP30729	Adp30729	Human	sec	185	17	50.0	1065	8	ADP31347	Adp31347	Human	sec
113	17	50.0	412	8	ADP57242	AdP57242	Plant	pol	186	17	50.0	1076	6	ABU81145	Abu81145	Human	PRO
114	17	50.0	420	8	ADP31273	Adp31273	Human	sec	187	17	50.0	1082	6	ABU66845	Abu66845	Human	PRO
115	17	50.0	420	8	ADP31274	Adp31274	Human	sec	188	17	50.0	1086	8	ADP30934	Adp30934	Human	sec
116	17	50.0	423	8	ADP31479	Adp31479	Human	sec	189	17	50.0	1092	8	ADP31358	Adp31358	Human	sec
117	17	50.0	483	8	ADP30852	Adp30852	Human	sec	190	17	50.0	1101	8	ADP31462	Adp31462	Human	sec
118	17	50.0	483	8	ADP30853	Adp30853	Human	sec	191	17	50.0	1110	8	ADP31430	Adp31430	Human	sec
119	17	50.0	510	8	ADP31049	Adp31049	Human	sec	192	17	50.0	1110	8	ADP31452	Adp31452	Human	sec
120	17	50.0	510	8	ADP31227	Adp31227	Human	sec	193	17	50.0	1128	8	ADP30483	Adp30483	Human	sec
121	17	50.0	528	8	ADP30529	Adp30529	Human	sec	194	17	50.0	1134	8	ADP30647	Adp30647	Human	sec
122	17	50.0	531	8	ADP31539	Adp31539	Human	sec	195	17	50.0	1134	8	ADP31537	Adp31537	Human	sec
123	17	50.0	531	8	ADP31540	Adp31540	Human	sec	196	17	50.0	1141	8	ADP30677	Adp30677	Human	sec
124	17	50.0	543	8	ADP30864	Adp30864	Human	sec	197	17	50.0	1170	8	ADP30919	Adp30919	Human	sec
125	17	50.0	552	8	ADP31019	Adp31019	Human	sec	198	17	50.0	1170	8	ADP30922	Adp30922	Human	sec
126	17	50.0	555	8	ADP31168	Adp31168	Human	sec	199	17	50.0	1173	8	ADP31155	Adp31155	Human	sec
127	17	50.0	559	8	ADP30716	Adp30716	Human	sec	200	17	50.0	1179	8	ADP30486	Adp30486	Human	sec
128	17	50.0	564	8	ADP30743	Adp30743	Human	sec	201	17	50.0	1189	8	ADP31043	Adp31043	Human	sec
129	17	50.0	567	8	ADP31565	Adp31565	Human	sec	202	17	50.0	1197	8	ADP31342	Adp31342	Human	sec
130	17	50.0	574	8	ADP31133	Adp31133	Human	sec	203	17	50.0	1205	8	ADP30935	Adp30935	Human	sec
131	17	50.0	592	8	ADP30917	Adp30917	Human	sec	204	17	50.0	1252	8	ADP30678	Adp30678	Human	sec
132	17	50.0	624	8	ADP31324	Adp31324	Human	sec	205	17	50.0	1358	8	ADP30995	Adp30995	Human	sec
133	17	50.0	624	8	ADP31325	Adp31325	Human	sec	206	17	50.0	1362	8	ADP31181	Adp31181	Human	sec
134	17	50.0	629	8	ADP31295	Adp31295	Human	sec	207	17	50.0	1383	8	ADP31091	Adp31091	Human	sec
135	17	50.0	645	8	ADP30858	Adp30858	Human	sec	208	17	50.0	1417	8	ADP31160	Adp31160	Human	sec
136	17	50.0	666	8	ADP31256	Adp31256	Human	sec	209	17	50.0	1420	8	ADP30944	Adp30944	Human	sec
137	17	50.0	681	8	ADP30598	Adp30598	Human	sec	210	17	50.0	1435	6	ABU09754	Abu09754	Human	hum
138	17	50.0	681	8	ADP30868	Adp30868	Human	sec	211	17	50.0	1435	6	ABU90136	Abu90136	Human	hum
139	17	50.0	702	8	ADP31518	Adp31518	Human	sec	212	17	50.0	1435	6	ABU964438	Abu964438	Human	hum
140	17	50.0	702	8	ADP31277	Adp31277	Human	sec	213	17	50.0	1435	6	ABU99047	Abu99047	Human	hum
141	17	50.0	702	8	ADP31278	Adp31278	Human	sec	214	17	50.0	1435	6	ABU98262	Abu98262	Human	hum
142	17	50.0	708	8	ADP31191	Adp31191	Human	sec	215	17	50.0	1435	6	ABU91968	Abu91968	Human	hum
143	17	50.0	711	8	ADP31535	Adp31535	Human	sec	216	17	50.0	1435	6	ABU08572	Abu08572	Human	hum
144	17	50.0	711	8	ADP31440	Adp31440	Human	sec	217	17	50.0	1435	6	ABU00411	Abu00411	Human	hum
145	17	50.0	714	8	ADP31561	Adp31561	Human	sec	218	17	50.0	1435	6	ABU88962	Abu88962	Human	hum
146	17	50.0	759	8	ADP31141	Adp31141	Human	sec	219	17	50.0	1435	6	ABU06458	Abu06458	Human	hum
147	17	50.0	768	8	ADP30912	Adp30912	Human	sec	220	17	50.0	1435	6	ABU95518	Abu95518	Human	hum
148	17	50.0	771	8	ADP30907	Adp30907	Human	sec	221	17	50.0	1435	6	ABU95208	Abu95208	Human	hum
149	17	50.0	774	8	ADP30885	Adp30885	Human	sec	222	17	50.0	1435	6	ABU90756	Abu90756	Human	hum
150	17	50.0	795	8	ADP31332	Adp31332	Human	sec	223	17	50.0	1435	6	ABU93918	Abu93918	Human	hum
151	17	50.0	799	8	ADP31261	Adp31261	Human	sec	224	17	50.0	1435	6	ABU86192	Abu86192	Human	hum
152	17	50.0	813	8	ADP30649	Adp30649	Human	sec	225	17	50.0	1435	6	ABU82047	Abu82047	Human	hum
153	17	50.0	834	8	ADP30738	Adp30738	Human	sec	226	17	50.0	1435	6	ABU94228	Abu94228	Human	hum
154	17	50.0	848	8	ADP31037	Adp31037	Human	sec	227	17	50.0	1435	6	ABU00101	Abu00101	Human	hum
155	17	50.0	849	8	ADP31182	Adp31182	Human	sec	228	17	50.0	1435	6	ABU07112	Abu07112	Human	hum
156	17	50.0	849	8	ADP31182	Adp31182	Human	sec	229	17	50.0	1435	6	ABU91353	Abu91353	Human	hum
157	17	50.0	849	8	ADP31275	Adp31275	Human	sec	230	17	50.0	1435	6	ABU90446	Abu90446	Human	hum
158	17	50.0	849	8	ADP31276	Adp31276	Human	sec	231	17	50.0	1435	6	ABU97037	Abu97037	Human	hum
159	17	50.0	861	8	ADP31340	Adp31340	Human	sec	232	17	50.0	1435	6	ABU05233	Abu05233	Human	hum
160	17	50.0	882	8	ADP30487	Adp30487	Human	sec	233	17	50.0	1464	8	ADP31437	Adp31437	Human	sec
161	17	50.0	895	8	ADP92573	Adp92573	Human	PRO	234	17	50.0	1472	8	ADP31611	Adp31611	Human	sec
162	17	50.0	900	6	ABU08402	Abu08402	Alpba-hel		235	17	50.0	1530	8	ADP31536	Adp31536	Human	sec
163	17	50.0	908	6	ADP31522	Adp31522	Human	sec	237	17	50.0	1550	8	ADP30567	Adp30567	Human	sec
164	17	50.0	921	8	ADP31444	Adp31444	Human	sec	238	17	50.0	1560	8	ADP31631	Adp31631	Human	sec
165	17	50.0	930	8	ADP31140	Adp31140	Human	sec	239	17	50.0	1614	8	ADP31529	Adp31529	Human	sec
166	17	50.0	933	8	ADP31542	Adp31542	Human	sec	240	17	50.0	1647	8	ADP30670	Adp30670	Human	sec
167	17	50.0	939	8	ADP31086	Adp31086	Human	sec	241	17	50.0	1662	8	ADP31419	Adp31419	Human	sec
168	17	50.0	939	8	ADP31541	Adp31541	Human	sec	242	17	50.0	1662	8	ADP31513	Adp31513	Human	sec
169	17	50.0	947	8	ADP30937	Adp30937	Human	sec	243	17	50.0	1700	3	AA18144	AA18144	Plaemoditu	















PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
(FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX PA  
XX PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Hhang MM, Kotlakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGB, Wu G, Zhang H;  
XX DR WPI; 2004-348438/32.  
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX PT genetic, bacterial and viral diseases.  
XX PS Claim 1; SEQ ID NO 2721; 428bp; English.  
XX CC The present invention relates to an isolated nucleic acid molecule  
XX CC encoding a polypeptide which is believed to be cytostatic,  
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX CC composition and methods are useful for diagnosing, preventing and  
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present  
XX CC sequence represents a human secreted protein. The present sequence is  
XX CC available on WIPWEB and is not in the specification.  
SQ Sequence 110 AA;

Query Match 52.9%; Score 18; DB 8; Length 110;  
Best Local Similarity 16.7%; Pred. No. 5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 93 CATTAATATTAAC 104

RESULT 3  
ADP31638  
ID ADP31638 standard; protein; 135 AA.  
XX AC ADP31638;  
XX DT 12-AUG-2004 (first entry)  
XX DE Human secreted protein SEQ ID #2405.  
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX KW cancer; inflammatory; immune; human secreted protein.  
XX OS Homo sapiens.  
XX PN WO2004035732-A2.  
XX PD 29-APR-2004.  
XX PF 28-AUG-2003; 2003WO-US026780.  
XX PR 29-AUG-2002; 2002US-0406576P.  
XX PR 29-AUG-2002; 2002US-0406579P.  
XX PR 29-AUG-2002; 2002US-0406585P.  
XX PR 29-AUG-2002; 2002US-0406588P.  
XX PR 29-AUG-2002; 2002US-0406608P.  
XX PR 29-AUG-2002; 2002US-0406611P.  
XX PR 29-AUG-2002; 2002US-0406612P.  
XX PR 29-AUG-2002; 2002US-0406616P.  
XX PR 29-AUG-2002; 2002US-0406642P.  
XX PR 29-AUG-2002; 2002US-0406646P.  
XX PR 29-AUG-2002; 2002US-0406653P.  
XX PR 29-AUG-2002; 2002US-0406655P.  
XX PR 29-AUG-2002; 2002US-0406666P.  
XX PR 17-SEP-2002; 2002US-0410946P.  
XX PR 17-SEP-2002; 2002US-0410947P.  
XX PR 17-SEP-2002; 2002US-0410948P.  
XX PR 17-SEP-2002; 2002US-0410949P.  
XX PR 17-SEP-2002; 2002US-0410953P.  
XX PR 17-SEP-2002; 2002US-0410957P.  
XX PR 17-SEP-2002; 2002US-0410958P.  
XX PR 17-SEP-2002; 2002US-0410959P.  
XX PR 17-SEP-2002; 2002US-0410960P.  
XX PR 17-SEP-2002; 2002US-0410961P.  
XX PR 17-SEP-2002; 2002US-0410962P.  
XX PR 17-SEP-2002; 2002US-0411019P.  
XX PR 17-SEP-2002; 2002US-0411022P.  
XX PR 17-SEP-2002; 2002US-0411023P.  
XX PR 17-SEP-2002; 2002US-0411024P.  
XX PR 17-SEP-2002; 2002US-0411028P.  
XX PR 17-SEP-2002; 2002US-0411032P.  
XX PR 17-SEP-2002; 2002US-0411035P.  
XX PR 17-SEP-2002; 2002US-0411037P.  
XX PR 17-SEP-2002; 2002US-0411041P.  
XX PR 17-SEP-2002; 2002US-0411045P.  
XX PR 17-SEP-2002; 2002US-0411046P.  
XX PR 17-SEP-2002; 2002US-0411048P.  
XX PR 17-SEP-2002; 2002US-0411052P.  
XX PR 17-SEP-2002; 2002US-0411055P.  
XX PR 17-SEP-2002; 2002US-0411073P.  
XX PR 17-SEP-2002; 2002US-0411082P.  
XX PR 17-SEP-2002; 2002US-0411101P.  
XX PR 17-SEP-2002; 2002US-0411111P.  
XX PR 18-APR-2003; 2003US-0463700P.  
XX PR 18-APR-2003; 2003US-0463708P.  
XX PR 18-APR-2003; 2003US-0463716P.  
XX PR 18-APR-2003; 2003US-0463732P.  
XX PR 02-MAY-2003; 2003US-0467199P.  
XX PR 02-MAY-2003; 2003US-0467201P.  
XX PR 02-MAY-2003; 2003US-0467203P.  
XX PR 02-MAY-2003; 2003US-0467230P.  
XX PR 19-MAY-2003; 2003US-0471306P.

```
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-047609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486911P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
DR
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3636; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPWEB and is not in the specification.
XX
XX Sequence 135 AA;
SQ
XX
XX Query Match 52.9%; Score 18; DB 8; Length 135;
XX Best Local Similarity 16.7%; Pred. No. 5.2e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
DB 89 CATTATATAAAC 100
XX
XX RESULT 4
XX ABO79745
XX ID ABO79745 standard; protein; 142 AA.
XX
XX ABO79745;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #11920.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
```

```
PR 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX N-PSDB; ABD13316.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 28491; 455bp; English.
XX
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using bioclip technology. Sequences ABO67826-
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX segdata.uspto.gov/sequence.html
XX
XX Sequence 142 AA;
SQ
XX
XX Query Match 52.9%; Score 18; DB 7; Length 142;
XX Best Local Similarity 16.7%; Pred. No. 5.2e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 4 CXXXXXXXXXXC 15
DB 88 CSASATSSATTC 99
XX
XX RESULT 5
XX ABO73284
XX ID ABO73284 standard; protein; 178 AA.
XX
XX ABO73284;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #5459.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX N-PSDB; ABD06855.
XX
```





```
XX Sequence 264 AA;
SQ
Query Match          52.9%; Score 18; DB 8; Length 264;
Best Local Similarity 16.7%; Pred. No. 6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy      4 CXXXXXXXXXXC 15
      109 CAAATATTTTC 120

RESULT 7
ADP31435
ID ADP31435 standard; protein; 270 AA.
XX
XX ADP31435;
AC
XX 12-AUG-2004 (first entry)
DT
XX
DE Human secreted protein SEQ ID #2202.
DE
XX
XX Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
```

```
PR 17-SEP-2002; 2002US-041101P.
PR 17-SEP-2002; 2002US-041111P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467203P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-048646P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
PI Halenbeck RP, Huang MW, Kothakota S, Halsehan L, Linemann T,
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX
XX PS Claim 1; SEQ ID NO 3433; 428pp; English.
XX
XX
XX CC The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPOMEB and is not in the specification.
XX
XX
XX SQ Sequence 270 AA;
XX
XX
XX Query Match          52.9%; Score 18; DB 8; Length 270;
XX Best Local Similarity 16.7%; Pred. No. 6e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy      4 CXXXXXXXXXXC 15
      179 CTATATTAAAC 190

RESULT 8
ABB65879
ID ABB65879 standard; protein; 307 AA.
XX
XX ABB65879;
AC
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 24429.
XX
```

KW	Drosophila; developmental biology; cell signalling; insecticide;
KV	pharmaceutical.
XX	
OS	Drosophila melanogaster.
PV	WO200171042-A2.
XX	
PD	27-SEP--2001.
XX	
PF	23-MAR-2001; 2001WO-US009231.
PR	23-MAR-2000; 2000US-0191637P.
PR	11-JUL-2000; 2000US-00614150.
XX	
PA	(PEKE ) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI, 2001-656860/75.
XX	N-PSTDB; ABL09982.
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
PT	
PS	Disclosure; SEQ ID NO 24429; 21pp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB822072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
CC	
CC	
XX	Sequence 307 AA;
XX	
SQ	
Query Match	52.9%; Score 18; DB 4; Length 307;
Best Local Similarity	16.7%; Pred.No. 6,2e-05;
Matches	2; Conservative 0; Mismatches 10; Indels 0; Gaps 0
OY	4 CXXXXXXXXXC 15
Db	44 CTTTTTTTTC 55
RESULT 9	
ADP31223	
ID	ADP31223 standard; protein; 357 AA.
XX	
AC	ADP31223;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Human secreted protein SEQ ID #1990.
XX	
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KV	cancer; inflammatory; immune; human secreted protein.
XX	
OS	Homo sapiens.
XX	
FN	WO2004035732-A2.
XX	
PD	29-APR--2004.
XX	
PF	28-AUG-2003; 2003WO-US026780.
XX	
PR	29-AUG-2002; 2002US-0406576P.
PR	29-AUG-2002; 2002US-0406579P.
PR	29-AUG-2002; 2002US-0406585P.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3221; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 357 AA;  
Query Match 52.9%; Score 18; DB 8; Length 357;  
Best Local Similarity 16.7%; Pred. No. 6.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Gy 4 CXXXXXXXXXC 15  
Db 175 CTAAATTATC 186  
RESULT 10  
ADP31132  
ID ADP31132 standard; protein; 615 AA.  
XX  
AC ADP31132;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1899.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411033P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411077P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Hsiehan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3130; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present  
XX sequence represents a human secreted protein. The present sequence is  
XX available on WIPOMEB and is not in the specification.  
SQ Sequence 615 AA;  
Query Match 52.9%; Score 18; DB 8; Length 615;  
Best Local Similarity 16.7%; Pred. No. 7.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
 DB 175 CATTATTTTAC 186

RESULT 11  
 AAU51580  
 ID AAU51580 standard; protein; 627 AA.  
 XX  
 AC AAU51580;  
 XX  
 DT 27-FEB-2002 (first entry)  
 XX  
 DE Propionibacterium acnes immunogenic protein #12476.  
 XX  
 KW SARPO syndrome; synovitis; acne; pustulosis; hypercosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 OS Propionibacterium acnes.  
 XX  
 PN WO200181581-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 PF 20-APR-2001; 2001WO-US012865.  
 XX  
 PR 21-APR-2000; 2000US-019047P.  
 PR 02-JUN-2000; 2000US-0208841P.  
 PR 07-JUL-2000; 2000US-0216747P.  
 XX  
 XX (CORI-) CORIXA CORP.  
 XX  
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A,  
 PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;  
 DR WPI: 2001-616774/71.  
 DR N-PSDB; AAS59551.  
 XX  
 PT Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris.  
 XX  
 PS Example 1; SEQ ID NO 12775; 1069pp; English.  
 CC Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SARPO syndrome (synovitis, acne,  
 CC pustulosis, hypercosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 CC Sequence 627 AA;  
 XX  
 SQ

Query Match 52.9%; Score 18; DB 4; Length 627;  
 Best Local Similarity 16.7%; Pred. No. 7.2e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
 DB 46 CATTTTTATSC 57

RESULT 12  
 ABM48099  
 ID ABM48099 standard; protein; 627 AA.  
 XX  
 AC ABM48099;  
 XX  
 DT 20-OCT-2003 (first entry)  
 XX  
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #12775.  
 XX  
 KW Acne vulgaris; antisporrothecic; dermatological; antibacterial;  
 KW immunostimulant; immune response; vaccine.  
 XX  
 OS Propionibacterium acnes.  
 XX  
 PN WO2003033515-A1.  
 XX  
 PD 24-APR-2003.  
 XX  
 PF 11-OCT-2002; 2002WO-US032727.  
 XX  
 PR 15-OCT-2001; 2001US-00978825.  
 XX  
 XX (CORI-) CORIXA CORP.  
 XX  
 PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL,  
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;  
 PI Barth B, Vallieve-Douglass J;  
 XX  
 DR WPI: 2003-381789/36.  
 DR N-PSDB; ACF64480.  
 XX  
 PT New Propionibacterium acnes polypeptides and polynucleotides encoding the  
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
 PT or for stimulating an immune response specific for a P. acnes protein.  
 XX  
 PS Example 1; SEQ ID NO 12775; 1481pp; English.  
 CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)  
 CC encoding a Propionibacterium acnes protein. The invention also relates to  
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to  
 CC immunogenic fragments of P. acnes polypeptides. The invention  
 CC additionally encompasses expression vectors and host cells comprising a  
 CC polynucleotide of the invention; antibodies against polypeptides of the  
 CC invention; fusion proteins comprising a polypeptide of the invention; a  
 CC method for stimulating an immune response specific for a P. acnes  
 CC polypeptide and an isolated T cell population comprising T cells prepared  
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,  
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or  
 CC antigen-presenting cells that express the polypeptide); a method and kit  
 CC for detecting or determining the presence or absence of P. acnes in a  
 CC patient; and a method for inhibiting the development of P. acnes in a  
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion  
 CC proteins, T cell populations or antigen-presenting cells that express the  
 CC polypeptides are useful for diagnosing, preventing or treating acne  
 CC vulgaris, or for stimulating an immune response specific for a P. acnes  
 CC protein. The polynucleotides can also be used as probes or primers for  
 CC nucleic acid hybridisation. The vaccine composition is useful for the  
 CC stimulation of an immune response against P. acnes, or for treating acne,  
 CC and the kit is useful for performing a diagnostic assay. The present  
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open  
 CC reading frame) contained within the P. acnes polynucleotides of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 CC Sequence 627 AA;  
 XX  
 SQ

Query Match 52.9%; Score 18; DB 6; Length 627;  
 Best Local Similarity 16.7%; Pred. No. 7.2e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 15  
 DB 46 CAATTTTATATSC 57

RESULT 13  
 ID ADP31142 standard; protein; 669 AA.  
 AC ADP31142;  
 DT 12-AUG-2004 (first entry)  
 XX Human secreted protein SEQ ID #1909.  
 DE  
 XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KW cancer; inflammatory; immune; human secreted protein.  
 OS Homo sapiens.  
 XX  
 PN WO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 XX  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406642P.  
 PR 29-AUG-2002; 2002US-0406646P.  
 PR 29-AUG-2002; 2002US-0406651P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 29-AUG-2002; 2002US-0406666P.  
 PR 17-SEP-2002; 2002US-0410946P.  
 PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410948P.  
 PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410958P.  
 PR 17-SEP-2002; 2002US-0410959P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410961P.  
 PR 17-SEP-2002; 2002US-0410962P.  
 PR 17-SEP-2002; 2002US-0411019P.  
 PR 17-SEP-2002; 2002US-0411022P.  
 PR 17-SEP-2002; 2002US-0411023P.  
 PR 17-SEP-2002; 2002US-0411024P.  
 PR 17-SEP-2002; 2002US-0411032P.  
 PR 17-SEP-2002; 2002US-0411035P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411041P.  
 PR 17-SEP-2002; 2002US-0411045P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411010P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR

PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471336P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 XX  
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX  
 PA William LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Huang MM, Kothakota S, Halahan L, Lammemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 DR WPI; 2004-348438/32.  
 XX  
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX  
 PS Claim 1; SEQ ID NO 3140; 428bp; English.  
 XX  
 CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMEB and is not in the specification.  
 XX  
 SQ Sequence 669 AA;  
 QY 52.9%; Score 18; DB 8; Length 669;  
 Best Local Similarity 16.7%; Pred. No. 7.3e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 15  
 DB 652 CAATTTTAAAC 663

RESULT 14  
 ID ADP31215 standard; protein; 711 AA.  
 AC ADP31215;  
 DT 12-AUG-2004 (first entry)  
 XX Human secreted protein SEQ ID #1982.  
 DE  
 XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KW cancer; inflammatory; immune; human secreted protein.  
 XX

OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485233P.  
PR 08-JUL-2003; 2003US-0485242P.  
PR 08-JUL-2003; 2003US-0485252P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3213; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
XX available on WIPOWEB and is not in the specification.  
SQ Sequence 711 AA;  
Query Match 52.9%; Score 18; DB 8; Length 711;  
Best Local Similarity 16.7%; Pred. No. 7.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXXXC 15  
Db 684 CAAAAAATTC 695  
RESULT 15  
ID ADP31196 standard; protein; 739 AA.  
XX  
XX ADP31196;  
AC  
XX 12-AUG-2004 (first entry)  
DT  
XX Human secreted protein SEQ ID #1963.  
DE  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485242P.  
PR 08-JUL-2003; 2003US-0485252P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
(FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PA  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haisman L, Linemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3194; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 739 AA;  
  
Query Match 52.9%; Score 18; DB 8; Length 739;  
Best Local Similarity 16.7%; Pred. No. 7,4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
Oy 4 CXXXXXXXXXXC 15  
Db 226 CAAAAAAAAAAC 237  
  
RESULT 16  
ADP31244  
ID ADP31244 standard; protein; 771 AA.  
XX  
XX ADP31244;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #2011.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406613P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.



PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411011P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485225P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

## (FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

## Claim 1; SEQ ID NO 3242; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMER and is not in the specification.

## SQ Sequence 771 AA;

Query Match 52.9%; Score 18; DB 8; Length 771;  
Best Local Similarity 16.7%; Pred. No. 7.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 662 CTTATTATTAAC 673

RESULT 17  
ADP31220  
ID ADP31220 standard; protein; 876 AA.  
XX  
AC ADP31220;  
XX

DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1987.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486860P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kotchakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
XX Claim 1; SEQ ID NO 3218; 428bp; English.  
XX  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
XX Sequence 876 AA;  
SQ  
Query Match 52.9%; Score 18; DB 8; Length 876;  
Best Local Similarity 16.7%; Pred. No. 7.7e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 4 CXXXXXXXXXC 15  
Db 28 CATATTAATC 39  
RESULT 18  
ADP31688  
ID ADP31688 standard; protein; 882 AA.  
XX  
XX ADP31688;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #2455.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX  
XX 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463715P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486860P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kotchakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3686; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.

XX  
XX Sequence 882 AA;

Query Match 52.9%; Score 18; DB 8; Length 882;  
Best Local Similarity 16.7%; Pred. No. 7.7e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 763 CTTATTATTAC 774

RESULT 19  
AAO14246  
ID AAO14246 standard; protein; 925 AA.

XX  
XX AAO14246;

XX  
XX 10-MAY-2002 (first entry)

XX  
XX Human presenilin enhancer protein pen-1B derived protein SEQ ID NO: 25.

XX  
XX Human; fruit fly; mouse; rat; cow; presenilin enhancer protein; pen;

XX  
XX Alzheimer's disease; pen-1; pen-1B; pen-2; Aph-2; amyloid beta.

XX  
XX Homo sapiens.

XX  
XX Synthetic.

XX  
XX WO200185912-A2.

XX  
XX 15-NOV-2001.

XX  
XX 03-MAY-2001; 2001WO-US014648.

XX  
XX 05-MAY-2000; 2000US-00568942.

XX  
XX (EXEL-) EXELIXIS INC.

XX  
XX Curtis DT, Francis GR, Ellis MC, Ruddy DA, Nicoll SM, McGrath GJ;

XX  
XX WPI, 2002-062245/08.

XX  
XX Presentin enhancer proteins and polynucleotides useful for modulating  
XX presentin function and screening for an agent that modulates the  
XX interaction of the protein to a binding target.

XX  
XX Disclosure: Page 67-70; 78pp; English.

XX  
XX The present invention relates to a method of detecting compounds capable  
CC of altering the interaction between a presenilin enhancer protein (such  
CC as pen-1, pen-1B, pen-2 and Aph-2) and presenilin. The inhibition of  
CC presentin activity causes the production of amyloid beta to be reduced  
CC and thus be used in the treatment of Alzheimer's disease. The present  
CC sequence is a presenilin enhancer protein described in the  
CC exemplification of the invention

XX  
XX Sequence 925 AA;

Query Match 52.9%; Score 18; DB 5; Length 925;  
Best Local Similarity 16.7%; Pred. No. 7.8e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 267 CATATTATTAAAC 278

RESULT 20  
ADP31517  
ID ADP31517 standard; protein; 1044 AA.

XX  
XX ADP31517;

XX  
XX 12-AUG-2004 (first entry)

XX  
XX Human secreted protein SEQ ID #2284.

XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.

XX  
XX Homo sapiens.

XX  
XX WO2004035732-A2.

XX  
XX 29-APR-2004.

XX  
XX 28-AUG-2003; 2003WO-US026780.

XX  
XX 29-AUG-2002; 2002US-0406576P.

XX  
XX 29-AUG-2002; 2002US-0406579P.

XX  
XX 29-AUG-2002; 2002US-0406585P.

XX  
XX 29-AUG-2002; 2002US-0406588P.

XX  
XX 29-AUG-2002; 2002US-0406608P.

XX  
XX 29-AUG-2002; 2002US-0406611P.

XX  
XX 29-AUG-2002; 2002US-0406613P.

XX  
XX 29-AUG-2002; 2002US-0406640P.

XX  
XX 29-AUG-2002; 2002US-0406642P.

XX  
XX 29-AUG-2002; 2002US-0406646P.

XX  
XX 29-AUG-2002; 2002US-0406653P.

XX  
XX 29-AUG-2002; 2002US-0406655P.

XX  
XX 29-AUG-2002; 2002US-0406666P.

XX  
XX 17-SEP-2002; 2002US-0410946P.

XX  
XX 17-SEP-2002; 2002US-0410947P.

XX  
XX 17-SEP-2002; 2002US-0410948P.

XX  
XX 17-SEP-2002; 2002US-0410949P.

XX  
XX 17-SEP-2002; 2002US-0410953P.

XX  
XX 17-SEP-2002; 2002US-0410957P.

XX  
XX 17-SEP-2002; 2002US-0410958P.

XX  
XX 17-SEP-2002; 2002US-0410961P.

XX  
XX 17-SEP-2002; 2002US-0410966P.

XX  
XX 17-SEP-2002; 2002US-0411019P.

XX  
XX 17-SEP-2002; 2002US-0411022P.

XX  
XX 17-SEP-2002; 2002US-0411023P.

XX  
XX 17-SEP-2002; 2002US-0411024P.

XX  
XX 17-SEP-2002; 2002US-0411032P.

XX  
XX 17-SEP-2002; 2002US-0411035P.

XX  
XX 17-SEP-2002; 2002US-0411037P.

XX  
XX 17-SEP-2002; 2002US-0411041P.

XX  
XX 17-SEP-2002; 2002US-0411045P.

XX  
XX 17-SEP-2002; 2002US-0411046P.

XX  
XX 17-SEP-2002; 2002US-0411048P.

XX  
XX 17-SEP-2002; 2002US-0411052P.

XX  
XX 17-SEP-2002; 2002US-0411055P.

XX  
XX 17-SEP-2002; 2002US-0411073P.

XX  
XX 17-SEP-2002; 2002US-0411082P.

XX  
XX 17-SEP-2002; 2002US-0411101P.

XX  
XX 18-APR-2003; 2003US-0463700P.

XX  
XX 18-APR-2003; 2003US-0463708P.

XX  
XX 18-APR-2003; 2003US-0463716P.

XX  
XX 18-APR-2003; 2003US-0463732P.

XX  
XX 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-047609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kotnakoca S, Halenhan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
DR WPI, 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 3515; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 1044 AA;  
XX  
Query Match 52.9%; Score 18; DB 8; Length 1044;  
Best Local Similarity 16.7%; Pred. No. 8e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
CY 4 CXXXXXXXXXXC 15  
DB 325 CATATAATTTC 336  
XX  
RESULT 21  
ADP31175  
ID ADP31175 standard; protein; 1086 AA.  
XX  
XX  
AC ADP31175;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1942.  
XX  
KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
XX  
PN WO2004035732-A2.  
XX

PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
XX  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411083P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-047609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 08-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 14-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX

Query Match	52.9%	Score 18	DB 8	Length 1086
Similarity	16.7%	Pred. No. 8e-05		
Best Local				
Matches	2	Conservative	10	Indels 0
				Gaps 0

RESULT 22  
ADP31508  
ID ADP31508 standard; protein; 1113 AA.

PR 17-SEP-2

XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases

XX	Claim 1; SEQ ID NO 3506; 428bp; English
PS	

The present invention relates to an isolated nucleic acid molecule encoding a polypeptide which is believed to be cytostatic, antiinflammatory, immunosuppressive, antibacterial and virucidal. The composition and methods are useful for diagnosing, preventing and treating diseases such as proliferative (e.g. cancer), inflammatory, immunologic, genetic, bacterial and viral diseases. The present sequence represents a human secreted protein. The present sequence is available on WIPWEB and is not in the specification.

XX SQ Sequence 1113 AA;  
Query Match 52.9%; Score 18; DB 8; Length 1113;  
Best Local Similarity 16.7%; Pred. No. 8.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
OY 4 CXXXXXXXXXXC 15  
DB 18 CATATTTTAAAC 29  
RESULT 23  
ADA15725  
ID ADA15725 standard; protein; 1128 AA.  
AC ADA15725;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE C. elegans neuromuscular junction GABA receptor complex subunit #6.  
XX  
KW Nematode;  
KM neuromuscular junction gamma-aminobutyric acid receptor complex;  
KM GABA receptor; parasitic plant pathogen; agricultural industry;  
KM crop protection; soil treatment.  
XX  
OS Caenorhabditis elegans.  
XX  
PN US2003065144-A1.  
XX  
PD 03-APR-2003.  
XX  
PF 24-MAY-2002; 2002US-00156240.  
XX  
PR 09-NOV-1998; 9AUS-01077227P.  
PR 08-NOV-1999; 99US-00436063.  
XX  
PA (UTAH ) UNIV UTAH RES FOUND.  
XX  
PI Bamber BA, Jorgensen EM;  
XX  
DR WPI: 2003-540802/51.  
DR N-PSDB; ADA15726.  
XX  
PT New nematode neuromuscular junction GABA receptor complex, useful for  
PT crop protection or soil treatment.  
XX  
PS Claim 21; Page 54-56; 84pp; English.  
XX  
CC The present invention relates to a nematode neuromuscular junction gamma-  
CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant  
CC pathogens which can cause major damage to crops in the agricultural  
CC industry. The nematode neuromuscular junction GABA receptor complex is  
CC useful for crop protection or soil treatment. The present sequence  
CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor  
CC complex subunit.  
XX  
SQ Sequence 1128 AA;  
Query Match 52.9%; Score 18; DB 6; Length 1128;  
Best Local Similarity 16.7%; Pred. No. 8.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
OY 4 CXXXXXXXXXXC 15  
DB 390 CAATATTTTTC 401  
RESULT 24  
ADP30741  
ID ADP30741 standard; protein; 1134 AA.  
XX

AC ADP30741;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1508.  
XX  
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406615P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406665P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463715P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.



PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1, SEQ ID NO 2922; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
SQ Sequence 1134 AA;  
  
Query Match 52.9%; Score 18; DB 8; Length 1134;  
Best Local Similarity 16.7%; Pred. No. 8.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 4 CXXXXXXXXXC 15  
DB 322 CTTAAAAATAC 333  
  
RESULT 26  
ADP31046  
ID ADP31046 standard; protein; 1168 AA.  
AC ADP31046;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1813.  
XX  
KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411083P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3044; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
SQ Sequence 1168 AA;  
  
Query Match 52.9%; Score 18; DB 8; Length 1168;  
Best Local Similarity 16.7%; Pred. No. 8.2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 4 CXXXXXXXXXC 15



DB 223 CAAATATTATAC 234

RESULT 27

ID ADP30993 standard; protein; 1191 AA.

XX ADP30993;

XX 12-AUG-2004 (first entry)

XX

DE Human secreted protein SEQ ID #1760.

XX

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

XX cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

XX

PN WO2004035732-A2.

XX

PD 29-APR-2004.

XX

PF 28-AUG-2003; 2003WO-US026780.

XX

XX 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406640P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410961P.

PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.

PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.

PR 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.

PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.

PR 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411011P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 08-JUL-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485225P.

PR 14-JUL-2003; 2003US-0486446P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

XX

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX

XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;

PI Halenbeck RF, Huang MW, Kotnakota S, Haldan L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;

XX

DR WPI; 2004-348438/32.

XX

XX New nucleic acid molecule for diagnosing, preventing or treating diseases

PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,

PT genetic, bacterial and viral diseases.

XX

PS Claim 1; SEQ ID NO 2991; 428bp; English.

XX

CC The present invention relates to an isolated nucleic acid molecule

CC encoding a polypeptide which is believed to be cytostatic,

CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and

CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is

CC available on WIPWEB and is not in the specification.

XX

SQ Sequence 1191 AA;

XX

Query Match 52.9%; Score 18; DB 8; Length 1191;

Best Local Similarity 16.7%; Pred. No. 8.2e-05;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 15

DB 821 CTTATATTATTC 832

RESULT 28

ADP31533

ID ADP31533 standard; protein; 1260 AA.

XX ADP31533;

XX 12-AUG-2004 (first entry)

XX

DE Human secreted protein SEQ ID #2300.

XX

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

XX cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

XX

PN WO2004035732-A2.

XX

PD 29-APR-2004.

XX

PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-041101P.  
PR 17-SEP-2002; 2002US-041111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485233P.  
PR 08-JUL-2003; 2003US-0485249P.  
PR 08-JUL-2003; 2003US-0485255P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halebbeck RF, Huang MM, Kotnakota S, Halsehan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX  
PS Claim 1, SEQ ID NO 3531; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
XX available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 1260 AA;  
XX  
Query Match 52.9%; Score 18; DB 8; Length 1260;  
Best Local Similarity 16.7%; Pred. No. 8.3e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
OY 4 CXXXXXXXXXC 15  
Db 757 CAATRAATATAC 768  
RESULT 29  
ADP30675  
ID ADP30675 standard; protein; 1289 AA.  
XX  
AC ADP30675;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1442.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Halsehan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX  
DR MPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX  
PS Claim 1, SEQ ID NO 2673; 428bp; English.  
XX  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present  
XX sequence represents a human secreted protein. The present sequence is  
XX available on WIPWEB and is not in the specification.  
SQ Sequence 1289 AA;

Query Match 52.9%; Score 18; DB 8; Length 1289;  
Best Local Similarity 16.7%; Pred. No. 8.3e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXXXC 15  
Db 1076 CATATATATTC 1087  
RESULT 30  
ADP31357  
ID ADP31357 standard; protein; 1437 AA.  
XX  
AC ADP31357;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2124.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; Inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406586P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411039P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486919P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493373P.  
PR 08-AUG-2003; 2003US-0493577P.  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3355; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 1437 AA;  
Query Match 52.9%; Score 18; DB 8; Length 1437;  
Best Local Similarity 16.7%; Pred. No. 8.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXXXC 15  
DB 206 CAATTAATATATC 217  
RESULT 31  
ADP31177  
ID ADP31177 standard; protein; 1454 AA.  
XX  
AC ADP31177;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #1944.  
DE  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.

XX OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.

PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.
XX		
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX		
PI	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;	
PI	Halenbeck RF, Huang MM, Kothakota S, Halshan L, Hinnemann T;	
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;	
XX		
DR	WPI: 2004-348438/32.	
XX		
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
PS	Claim 1; SEQ ID NO 3175; 428bp; English.	
XX		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
CC	available on WIPOMEB and is not in the specification.	
XX		
SQ	Sequence 1454 AA;	
QY	4 CXXXXXXXXXC 15	
DB	1094 CATTATATATAC 1105	
RESULT 32		
ADP30557		
ID	ADP30557 standard; protein; 1480 AA.	
XX		
AC	ADP30557;	
XX		
DT	12-AUG-2004 (first entry)	
DX		
DE	Human secreted protein SEQ ID #1324.	
XX		
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;	
KW	cancer; inflammatory; immune; human secreted protein.	
XX		
OS	Homo sapiens.	
XX		
PN	WO2004035732-A2.	
XX		
PD	29-APR-2004.	
XX		
PF	28-AUG-2003; 2003WO-US026780.	
XX		
PR	29-AUG-2002;	2002US-0406576P.
PR	29-AUG-2002;	2002US-0406579P.
PR	29-AUG-2002;	2002US-0406585P.
PR	29-AUG-2002;	2002US-0406588P.
PR	29-AUG-2002;	2002US-0406608P.
PR	29-AUG-2002;	2002US-0406611P.
PR	29-AUG-2002;	2002US-0406612P.
PR	29-AUG-2002;	2002US-0406616P.
PR	29-AUG-2002;	2002US-0406640P.
PR	29-AUG-2002;	2002US-0406643P.
PR	29-AUG-2002;	2002US-0406646P.
PR	29-AUG-2002;	2002US-0406653P.
PR	29-AUG-2002;	2002US-0406655P.

CC	antiinflammatory, immunosuppressive, antibacterial and antiviral diseases.
PS	Claim 1; SEQ ID NO 2555; 428bp; English.
XX	New nucleic acid molecule for diagnosing, preventing or treating diseases
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
FT	genetic, bacterial and viral diseases.
XX	
XX	
DR	WPI; 2004-348438/32.
XX	
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.
XX	
P1	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
P1	Halemebeck RT, Huang MM, Kotchakota S, Hashish L, Linnemann T;
P1	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX	
PR	29-AUG-2002; 2002US-0406666P.
PR	17-SEP-2002; 2002US-0410946P.
PR	17-SEP-2002; 2002US-0410947P.
PR	17-SEP-2002; 2002US-0410948P.
PR	17-SEP-2002; 2002US-0410949P.
PR	17-SEP-2002; 2002US-0410953P.
PR	17-SEP-2002; 2002US-0410957P.
PR	17-SEP-2002; 2002US-0410958P.
PR	17-SEP-2002; 2002US-0410959P.
PR	17-SEP-2002; 2002US-0410960P.
PR	17-SEP-2002; 2002US-0410961P.
PR	17-SEP-2002; 2002US-0410962P.
PR	17-SEP-2002; 2002US-0411019P.
PR	17-SEP-2002; 2002US-0411022P.
PR	17-SEP-2002; 2002US-0411023P.
PR	17-SEP-2002; 2002US-0411024P.
PR	17-SEP-2002; 2002US-0411032P.
PR	17-SEP-2002; 2002US-0411035P.
PR	17-SEP-2002; 2002US-0411037P.
PR	17-SEP-2002; 2002US-0411052P.
PR	17-SEP-2002; 2002US-0411055P.
PR	17-SEP-2002; 2002US-0411073P.
PR	17-SEP-2002; 2002US-0411082P.
PR	17-SEP-2002; 2002US-0411101P.
PR	17-SEP-2002; 2002US-0411111P.
PR	18-APR-2003; 2003US-0463708P.
PR	18-APR-2003; 2003US-0463708P.
PR	18-APR-2003; 2003US-0463716P.
PR	18-APR-2003; 2003US-0463732P.
PR	02-MAY-2003; 2003US-0467199P.
PR	02-MAY-2003; 2003US-0467201P.
PR	02-MAY-2003; 2003US-0467203P.
PR	02-MAY-2003; 2003US-0467230P.
PR	19-MAY-2003; 2003US-0471306P.
PR	19-MAY-2003; 2003US-0471336P.
PR	22-MAY-2003; 2003US-0472420P.
PR	22-MAY-2003; 2003US-0472430P.
PR	09-JUN-2003; 2003US-0476609P.
PR	09-JUN-2003; 2003US-0476641P.
PR	08-JUL-2003; 2003US-0485218P.
PR	08-JUL-2003; 2003US-0485223P.
PR	08-JUL-2003; 2003US-0485224P.
PR	08-JUL-2003; 2003US-0485325P.
PR	14-JUL-2003; 2003US-0486446P.
PR	14-JUL-2003; 2003US-0486480P.
PR	15-JUL-2003; 2003US-0486891P.
PR	15-JUL-2003; 2003US-0486960P.
PR	08-AUG-2003; 2003US-0493341P.
PR	08-AUG-2003; 2003US-0493370P.
PR	08-AUG-2003; 2003US-0493573P.
PR	08-AUG-2003; 2003US-0493577P.

CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.

XX  
SQ Sequence 1480 AA;

Query Match 52.9%; Score 18; DB 8; Length 1480;  
Best Local Similarity 16.7%; Pred. No. 8.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 1286 CTPAAATATTAAC 1297

#### RESULT 33

ID ABB09437 standard; protein; 1588 AA.

XX ABB09437;

XX 01-JUL-2002 (first entry)

XX H. influenzae DXR related polypeptide sequence.

XX DXR; reductoisomerase; enzyme; non-mevalonate isoprenoid; menaquinone;

XX ubiquinone; virulence; ear infection; conjunctivitis; meningitis;

XX pneumonia; conjunctivitis; bacteraemia; sinusitis; pleural empyema;

XX endocarditis; epiglottitis.

XX Haemophilus influenzae.

XX Key Location/Qualifiers

XX Region 241..1431

XX WO200211673-A2.

XX 14-FEB-2002.

XX 09-AUG-2001; 2001WO-US024950.

XX 09-AUG-2000; 2000US-0223909P.

XX (SMK) SMITHKLINE BEECHAM CORP.

XX (SMK) SMITHKLINE BEECHAM PLC.

XX Jaworski DD, Payne DJ, Slater-Radosci CE, Yan K;

XX WPI; 2002-241698/29.

XX Modulating Haemophilus influenzae DXR reductoisomerase enzyme activity,

XX useful for treating mammals or tissues infected with H. influenzae (e.g. a

XX ear infections or pneumonia) by contacting the enzyme with a modulator of

XX its activity.

XX Disclosure; Page 40-44; 44pp; English.

XX The invention relates to modulating an activity of a DXR reductoisomerase

XX enzyme of Haemophilus influenzae, comprising contacting the enzyme with a

XX compound that modulates non-mevalonate isoprenoid biosynthesis -

XX synthesis of menaquinone or ubiquinone. Compounds of the invention act as

XX virucides. The method is useful for treating a mammal or mammalian tissue

XX infected with H. influenzae having DXR reductoisomerase enzyme, e.g. a

XX human or a domestic animal. In particular, the method is useful for

XX treating ear infections, conjunctivitis, meningitis, pneumonia,

XX conjunctivitis, bacteraemia, sinusitis, pleural empyema, endocarditis and

XX epiglottitis. The current sequence represents a H. influenzae DXR

CC reductoisomerase enzyme related polypeptide sequence. Note: The current  
CC sequence contains within it the amino acid sequence given in record  
CC ABB09436 (DXR enzyme), but this is broken up by a large insertion that  
CC appears to be accidentally inserted into the sequence, consisting the DXR  
CC encoding DNA sequence represented as an amino acid sequence in three  
CC letter code

XX  
SQ Sequence 1588 AA;

Query Match 52.9%; Score 18; DB 5; Length 1588;  
Best Local Similarity 16.7%; Pred. No. 8.7e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 431 CTPAAATTTTAC 442

#### RESULT 34

ID ADP30660 standard; protein; 1617 AA.

XX ADP30660;

XX 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #1427.

XX Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;

XX cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX WO2004035732-A2.

XX 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411101P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471336P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485225P.  
 PR 14-JUL-2003; 2003US-0486448P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 (FIVE-) FIVE PRIME THERAPEUTICS INC.

PA William LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Huang MM, Kotchakota S, Haispan L, Linnemann T;  
 PI Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;  
 DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.

PS Claim 1; SEQ ID NO 2658; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cyrostatic.  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMEB and is not in the specification.

XX Sequence 1617 AA;

Query Match 52.9%; Score 18; DB 8; Length 1617;  
 Best Local Similarity 16.7%; Pred. No. 8.7e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
 DB 71 CTATATTAATTC 82

RESULT 35  
 ADA15715  
 ID ADA15715 standard; protein; 1652 AA.  
 XX

AC ADA15715;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE C. elegans neuromuscular junction GABA receptor complex subunit #1.  
 XX  
 KW Nematode;  
 KW neuromuscular junction gamma-aminobutyric acid receptor complex;  
 KW GABA receptor; parasitic plant pathogen; agricultural industry;  
 KW crop protection; soil treatment.  
 XX  
 OS Caenorhabditis elegans.  
 XX  
 PN US2003065144-A1.  
 XX  
 PD 03-APR-2003.  
 XX  
 PF 24-MAY-2002; 2002US-00156240.  
 XX  
 PR 09-NOV-1998; 98US-0107727P.  
 PR 08-NOV-1999; 99US-00436063.  
 XX  
 PA (UTAH) UNITV UTAH RES FOUND.  
 XX  
 PI Bamber BA, Jorgensen EM;  
 XX  
 DR WPI; 2003-540802/51.  
 XX  
 PT New nematode neuromuscular junction GABA receptor complex, useful for  
 PT crop protection or soil treatment.  
 XX  
 PS Claim 21; Page 20-24; 84bp; English.

XX The present invention relates to a nematode neuromuscular junction gamma-  
 CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant  
 CC pathogens which can cause major damage to crops in the agricultural  
 CC industry. The nematode neuromuscular junction GABA receptor complex is  
 CC useful for crop protection or soil treatment. The present sequence  
 CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor  
 CC complex subunit.

XX Sequence 1652 AA;

Query Match 52.9%; Score 18; DB 6; Length 1652;  
 Best Local Similarity 16.7%; Pred. No. 8.8e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
 DB 891 CAATATTTTTC 902

RESULT 36  
 ADP30654  
 ID ADP30654 standard; protein; 1725 AA.

XX ADP30654;

XX 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #1421.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN WO2004035732-A2.

PD 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

XX

PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485255P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Halshah L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2652; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 1725 AA;  
XX  
Query Match 52.9%; Score 18; DB 8; Length 1725;  
Best Local Similarity 16.7%; Pred. No. 8.9e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 4 CXXXXXXXXXXC 15  
Db 828 CTAATTTAAAC 839  
RESULT 37  
ADP30642  
ID ADP30642 standard; protein; 1833 AA.  
XX  
XX ADP30642;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1409.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
PN  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0405888P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.



PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411011P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 14-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kohakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2640; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
XX Sequence 1833 AA;  
SQ

Query Match

52.9%; Score 18; DB 8; Length 1833;

Best Local Similarity 16.7%; Pred. No. 9e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXXC 15  
Db 462 CTATTTTAAAC 473  
RESULT 38  
ADP31178  
ID ADP31178 standard; protein; 2088 AA.  
XX  
XX ADP31178;  
AC  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1945.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411101P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 18-APR-2003; 2003US-0463700P.  
XX 18-APR-2003; 2003US-0463708P.

PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US-0463732E.
PR	02-MAY-2003;	2003US-0467195P.
PR	02-MAY-2003;	2003US-0467201P.
PR	02-MAY-2003;	2003US-0467203P.
PR	02-MAY-2003;	2003US-0467230P.
PR	19-MAY-2003;	2003US-0471306P.
PR	19-MAY-2003;	2003US-0471336P.
PR	22-MAY-2003;	2003US-0472420P.
PR	22-MAY-2003;	2003US-0472430P.
PR	09-JUN-2003;	2003US-0476609P.
PR	09-JUN-2003;	2003US-0476641P.
PR	08-JUL-2003;	2003US-0485218P.
PR	08-JUL-2003;	2003US-0485223P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486446P.
PR	14-JUL-2003;	2003US-0486480P.
PR	15-JUL-2003;	2003US-0486891P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.
XX	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
PA		
XX		
PI	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;	
PI	Halenbeck RF, Huang MM, Koethakota S, Halpan L, Linemann T;	
PI	Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;	
DR	WPI; 2004-348438/32.	
XX		
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
ES	Claim 1; SEQ ID NO 3176; 428bp; English.	
XX		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
XX	available on WIPOMEB and is not in the specification.	
SQ	Sequence 2088 AA;	
Query Match	52.9%;	Score 18; DB 8; Length 2088;
Best Local Similarity	16.7%;	Pred. No. 9.2e-05;
Matches	2; Conservative	0; Mismatches 10; Indels 0; Gaps 0.
CY	4 CXXXXXXXXXC 15	
DB	1094 CATTTATATAAC 1105	
RESULT 39		
ID	ADP31327	
ID	ADP31327 standard; protein; 2127 AA.	
AC	ADP31327;	
XX		
DT	12-AUG-2004 (first entry)	
DE	Human secreted protein SEQ ID #2094.	
XX		
KM	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;	
KW	cancer; inflammatory; immune; human secreted protein.	
OS	Homo sapiens.	

[illegible]

PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D;  
PI Halebbeck RF, Huang MM, Kochakota S, Haisan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1, SEQ ID NO 3325; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 2127 AA;  
XX  
XX Query Match 52.9%; Score 18; DB 8; Length 2127;  
Best Local Similarity 16.7%; Pred. No. 9.2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
XX  
QY 4 CXXXXXXXXXX 15  
DB 1464 CTTATTAAAAAC 1475  
XX  
XX RESULT 40  
ADP66690  
ID ADP66690 standard; protein; 2484 AA.  
XX  
XX ADP66690;  
XX  
XX 26-AUG-2004 (first entry)  
XX  
XX Human mismatch repair protein MLH1.  
XX  
XX PMS1; PMS2; PMSR3; PMSR2; PMSR6; MLH1; GTBP; MSH3; MSH1; PMSR;  
XX immunoglobulin; mismatch repair protein; human.  
XX  
XX Homo sapiens.  
XX  
XX WO2004046330-A2.  
XX  
XX 03-JUN-2004.  
XX  
XX 14-NOV-2003; 2003WO-US036702.  
XX  
XX 15-NOV-2002; 2002US-0427165P.  
XX  
XX 10-SEP-2003; 2003US-0501650P.  
XX  
XX (MORP-) MORPHOTEK INC.  
XX  
XX Graess L, Liang S, Nicolaides NE, Sass PM;  
XX WPI; 2004-440979/41.  
XX  
XX Producing mammalian expression cells producing high-affinity antibodies  
PT from immunized immunoglobulin-producing cells, by combining cells with  
PT antigen, forming parental hybridoma cells and hypermutated hybridoma  
PT cells, using myeloma cells.  
XX  
XX Claim 135; SEQ ID NO 12; 213bp; English.

XX  
XX The invention relates to producing mammalian expression cells e.g. in  
CC hybridoma cells producing high-affinity and high titer antibodies from in  
CC vitro immunized immunoglobulin-producing cells. The method involves:  
CC combining donor cells comprising immunoglobulin (Ig)-producing cells with  
CC an immunogenic antigen in vitro, fusing the Ig-producing cells with  
CC myeloma cells to form parental hybridoma cells, where the hybridoma cells  
CC express a dominant negative allele of a mismatch repair gene, incubating  
CC the parental hybridoma cells to allow for mutagenesis, thus forming  
CC hypermutated hybridoma cells, performing a screen for binding of  
CC antibodies to antigen for antibodies produced from the hypermutated  
CC hybridoma cells, and selecting hypermutated hybridoma cells that produce  
CC antibodies with greater affinity for the antigen than antibodies produced  
CC by the parental hybridoma cells, thus producing hybridoma cells producing  
CC high-affinity antibodies. In the method, the dominant negative allele of  
CC a mismatch repair gene comprises a dominant negative allele of a gene  
CC chosen from PMS2, PMS1, PMSR3, PMSR2, PMSR6, MLH1, GTBP, MSH3, MSH1, or  
CC MSH1, and homologues of PMSR genes. The method is useful for producing  
CC mammalian expression cells e.g., hybridoma cells producing high-affinity  
CC and high titer antibodies from in vitro immunized immunoglobulin-  
CC producing cells. The present sequence represents a human mismatch repair  
CC protein MLH1.  
XX  
SQ Sequence 2484 AA;  
XX  
XX Query Match 52.9%; Score 18; DB 8; Length 2484;  
Best Local Similarity 16.7%; Pred. No. 9.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
XX  
QY 4 CXXXXXXXXXX 15  
DB 515 CTTTAAAAAATC 526  
XX  
XX RESULT 41  
ADA15721  
ID ADA15721 standard; protein; 2508 AA.  
XX  
XX ADA15721;  
XX  
XX 06-NOV-2003 (first entry)  
XX  
XX C. elegans neuromuscular junction GABA receptor complex subunit #4.  
XX  
XX Nematode;  
XX neuromuscular junction gamma-aminobutyric acid receptor complex;  
XX GABA receptor; parasitic plant pathogen; agricultural industry;  
XX crop protection; soil treatment.  
XX  
XX Caenorhabditis elegans.  
XX  
XX US2003065144-A1.  
XX  
XX 03-APR-2003.  
XX  
XX 24-MAY-2002; 2002US-00156240.  
XX  
XX 09-NOV-1998; 98US-0107727P.  
XX  
XX 08-NOV-1999; 99US-00436063.  
XX  
XX (UTAH ) UNIV UTAH RES FOUND.  
XX  
XX Bamber BA, Jorgensen EM;  
XX WPI; 2003-540802/51.  
XX  
XX N-PSDB; ADA15722.  
XX  
XX New nematode neuromuscular junction GABA receptor complex, useful for  
PT crop protection or soil treatment.  
XX  
XX Claim 21; Page 38-44; 84bp; English.  
XX  
XX The present invention relates to a nematode neuromuscular junction gamma-

CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant  
CC pathogens which can cause major damage to crops in the agricultural  
CC industry. The nematode neuromuscular junction GABA receptor complex is  
CC useful for crop protection or soil treatment. The present sequence  
CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor  
CC complex subunit.

XX  
SQ Sequence 2508 AA;

Query Match 52.9%; Score 18; DB 6; Length 2508;  
Best Local Similarity 16.7%; Pred. No. 9.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15  
Db 1775 CAAATATTTTTC 1786

#### RESULT 42

ID ADA15717 standard; protein; 2544 AA.

XX ADA15717;

AC 06-NOV-2003 (first entry)

XX C. elegans neuromuscular junction GABA receptor complex subunit #2.

XX Nematode;

KW neuromuscular junction gamma-aminobutyric acid receptor complex;

KM GABA receptor; parasitic plant pathogen; agricultural industry;

KW crop protection; soil treatment.

OS Caenorhabditis elegans.

XX US2003065144-A1.

XX 03-APR-2003.

XX 24-MAY-2002; 2002US-00156240.

XX 09-NOV-1998; 98US-0107727P.

XX 08-NOV-1999; 99US-00436063.

XX (UTAH ) UNIV UTAH RES FOUND.

XX Bamber BA, Jorgensen EM;

XX WPI; 2003-540802/51.

XX N-PSDB; ADA15718.

XX New nematode neuromuscular junction GABA receptor complex, useful for

XX crop protection or soil treatment.

XX Claim 21; Page 25-31; 84pp; English.

XX The present invention relates to a nematode neuromuscular junction gamma-

XX aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant

XX pathogens which can cause major damage to crops in the agricultural

XX industry. The nematode neuromuscular junction GABA receptor complex is

XX useful for crop protection or soil treatment. The present sequence

XX represents a Caenorhabditis elegans neuromuscular junction GABA receptor

XX complex subunit.

XX Sequence 2544 AA;

RESULT 43  
ADA15723  
ID ADA15723 standard; protein; 2601 AA.

XX ADA15723;

AC 06-NOV-2003 (first entry)

XX C. elegans neuromuscular junction GABA receptor complex subunit #5.

XX Nematode;

KW neuromuscular junction gamma-aminobutyric acid receptor complex;

KM GABA receptor; parasitic plant pathogen; agricultural industry;

KW crop protection; soil treatment.

OS Caenorhabditis elegans.

XX US2003065144-A1.

XX 03-APR-2003.

XX 24-MAY-2002; 2002US-00156240.

XX 09-NOV-1998; 98US-0107727P.

XX 08-NOV-1999; 99US-00436063.

XX (UTAH ) UNIV UTAH RES FOUND.

XX Bamber BA, Jorgensen EM;

XX WPI; 2003-540802/51.

XX N-PSDB; ADA15724.

XX New nematode neuromuscular junction GABA receptor complex, useful for

XX crop protection or soil treatment.

XX Claim 21; Page 46-52; 84pp; English.

XX The present invention relates to a nematode neuromuscular junction gamma-

XX aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant

XX pathogens which can cause major damage to crops in the agricultural

XX industry. The nematode neuromuscular junction GABA receptor complex is

XX useful for crop protection or soil treatment. The present sequence

XX represents a Caenorhabditis elegans neuromuscular junction GABA receptor

XX complex subunit.

XX Sequence 2601 AA;

Query Match 52.9%; Score 18; DB 6; Length 2601;  
Best Local Similarity 16.7%; Pred. No. 9.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15  
Db 1251 CAAATATTAATTC 1262

RESULT 44  
ADP31299  
ID ADP31299 standard; protein; 2833 AA.  
XX ADP31299;  
XX 12-AUG-2004 (first entry)  
XX Human secreted protein SEQ ID #2066.  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX Homo sapiens.

[illegible]

PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.
XX		
PA	(FIVE-)	FIVE PRIME THERAPEUTICS INC.
XX		
PI	Williams LT, Chu K, Lee E, Heattir K, Beaurang PA, Behrens D;	
PI	Haltenbeck RF, Huang MM, Kochakota S, Haishan L, Linnewann T;	
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;	
XX		
DR	WPI; 2004-348438/32.	
XX		
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
PS	Claim 1; SEQ ID NO 3297; 428bp; English.	
XX		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cyrostatic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
CC	available on WIPOMEB and is not in the specification.	
SQ	Sequence 2833 AA;	
Query Match	52.9%; Score 18; DB 8; Length 2833;	
Best Local Similarity	16.7%; Pred. No. 9.8e-05;	
Matches	2; Conservative 0; Mismatches 10; Indels 0; Gaps 0.	
OY	4 CXXXXXXXXXXC 15	
DB	1876 CATAAATTATTC 1887	
RESULT 45		
ADP30572		
ID	ADP30572 standard; protein, 2835 AA.	
XX		
AC	ADP30572;	
XX		
DT	12-AUG-2004 (first entry)	
XX		
DE	Human secreted protein SEQ ID #1339.	
XX		
KW	Cyostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;	
KW	cancer; inflammatory; immune; human secreted protein.	
XX		
OS	Homo sapiens.	
XX		
WO	WO2004035732-A2.	
XX		
PD	29-APR-2004.	
XX		
PF	28-AUG-2003; 2003WO-US026780.	
XX		
PR	29-AUG-2002; 2002US-0406576P.	
PR	29-AUG-2002; 2002US-0406579P.	
PR	29-AUG-2002; 2002US-0406585P.	
PR	29-AUG-2002; 2002US-0406588P.	
PR	29-AUG-2002; 2002US-0406608P.	
PR	29-AUG-2002; 2002US-0406611P.	
PR	29-AUG-2002; 2002US-0406612P.	
PR	29-AUG-2002; 2002US-0406616P.	
PR	29-AUG-2002; 2002US-0406640P.	
PR	29-AUG-2002; 2002US-0406642P.	
PR	29-AUG-2002; 2002US-0406646P.	
PR	29-AUG-2002; 2002US-0406653P.	
PR	29-AUG-2002; 2002US-0406655P.	
PR	29-AUG-2002; 2002US-0406666P.	
PR	17-SEP-2002; 2002US-0410946P.	

PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486860P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halesbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,  
XX  
XX WPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 2570; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
XX  
SQ Sequence 2835 AA;  
Query Match 52.9%; Score 18; DB 8; Length 2835;  
Best Local Similarity 16.7%; Pred. No. 9.8e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 4 CXXXXXXXXXXC 15  
Db 884 CAAAAATTATTC 895  
RESULT 46  
ADP30667  
ID ADP30667 standard; protein; 3411 AA.  
XX  
XX ADP30667;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
DT XX  
XX  
DE Human secreted protein SEQ ID #1434.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
OS  
XX  
PN MO2004035732-A2.  
PD  
XX  
XX 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

## (FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halsebeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX MPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

PS Claim 1: SEQ ID NO 2665; 428bp; English.

CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytosolic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.

XX Sequence 3411 AA;

QY Query Match 52.9%; Score 18; DB 8; Length 3411;  
Best Local Similarity 16.7%; Pred. No. 0.0001;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Db 4 CXXXXXXXXXXC 15  
2826 CAATTTATTTTC 2837

## RESULT 47

ABU88256  
ID ABU88256 standard; protein; 4440 AA.  
AC ABU88256;  
XX  
DT 07-JUL-2003 (first entry)

XX Novel human secreted and transmembrane PRO polypeptide #4.  
DE Human; secreted and transmembrane protein; PRO; gene therapy;  
XX tumour necrosis factor-alpha release; TNF-alpha release;  
XX chondrocyte proliferation; chondrocyte differentiation; tumour;  
XX adrenal tumour; lung tumour; colon tumour; breast tumour;  
XX prostate tumour; rectal tumour; cervical tumour; liver tumour.  
OS Homo sapiens.  
XX  
XX US2003032127-A1.  
XX  
PD 13-FEB-2003.  
XX  
PF 26-JUN-2002; 2002US-00183012.  
XX  
XX 18-SEP-1997; 97US-0059263P.  
XX 18-SEP-1997; 97US-0059266P.  
XX 17-OCT-1997; 97US-0062250P.  
XX 21-OCT-1997; 97US-0063486P.  
XX 24-OCT-1997; 97US-0063120P.  
XX 24-OCT-1997; 97US-0063121P.  
XX 28-OCT-1997; 97US-0063540P.  
XX 28-OCT-1997; 97US-0063541P.  
XX 28-OCT-1997; 97US-0063544P.  
XX 29-OCT-1997; 97US-0063734P.  
XX 31-OCT-1997; 97US-0063870P.  
XX 31-OCT-1997; 97US-0064103P.  
XX 13-NOV-1997; 97US-0065119P.  
XX 24-NOV-1997; 97US-0065120P.  
XX 24-NOV-1997; 97US-0066466P.  
XX 11-DEC-1997; 97US-0069335P.  
XX 12-DEC-1997; 97US-0069425P.  
XX 17-DEC-1997; 97US-0069870P.  
XX 18-DEC-1997; 97US-0068017P.  
XX 10-MAR-1998; 98US-0077450P.  
XX 11-MAR-1998; 98US-0077632P.  
XX 11-MAR-1998; 98US-0077649P.  
XX 20-MAR-1998; 98US-0078886P.  
XX 20-MAR-1998; 98US-0078939P.  
XX 27-MAR-1998; 98US-0079664P.  
XX 27-MAR-1998; 98US-0079786P.  
XX 31-MAR-1998; 98US-0080107P.  
XX 31-MAR-1998; 98US-0080194P.  
XX 01-APR-1998; 98US-0080327P.  
XX 01-APR-1998; 98US-0080333P.  
XX 08-APR-1998; 98US-0081049P.  
XX 08-APR-1998; 98US-0081070P.  
XX 09-APR-1998; 98US-0081195P.  
XX 15-APR-1998; 98US-0081838P.  
XX 21-APR-1998; 98US-0082568P.  
XX 21-APR-1998; 98US-0082569P.  
XX 22-APR-1998; 98US-0082704P.  
XX 22-APR-1998; 98US-0083797P.  
XX 28-APR-1998; 98US-0083722P.  
XX 29-APR-1998; 98US-0083495P.  
XX 29-APR-1998; 98US-0083496P.  
XX 29-APR-1998; 98US-0083499P.  
XX 29-APR-1998; 98US-0083559P.  
XX 05-MAY-1998; 98US-0084366P.  
XX 06-MAY-1998; 98US-0084414P.  
XX 07-MAY-1998; 98US-0084639P.  
XX 07-MAY-1998; 98US-0084640P.  
XX 07-MAY-1998; 98US-0084643P.  
XX 15-MAY-1998; 98US-0085579P.  
XX 15-MAY-1998; 98US-0085580P.  
XX 15-MAY-1998; 98US-0085582P.  
XX 15-MAY-1998; 98US-0085700P.  
XX 18-MAY-1998; 98US-0086023P.  
XX 22-MAY-1998; 98US-0086392P.

PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-008812P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-008811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 22-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.

PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-01019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 18-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.

Query Match 52.9%; Score 18; DB 6; Length 4440;  
Best Local Similarity 16.7%; Pred. No. 0.00011;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

DB 3858 CATTTTAATAC 3869

RESULT 48  
ABU90135  
ID ABU90135 standard; protein; 4440 AA.

AC ABU90135;

DT 11-AUG-2003 (first entry)

XX Novel human secreted and transmembrane PRO protein #4.

XX Human; gene therapy; tissue typing; tumour; chondrocyte proliferation;  
KW chondrocyte differentiation; tumour necrosis factor-alpha release;  
XX affinity purification.

XX



OS Homo sapiens.  
XX  
XX US2003036147-A1.  
XX  
PD 20-FEB-2003.  
XX  
XX 02-JUL-2002; 2002US-00187741.  
XX  
PR 18-SRP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069670P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 09-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088555P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P\*.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090680P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095988P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096577P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.

PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0098602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0101933P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101016P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-010230P.  
PR 29-SEP-1998; 98US-010231P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102865P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-010349P.  
PR 07-OCT-1998; 98US-00168978.

Query Match 52.9%; Score 18; DB 6; Length 4440;  
Best Local Similarity 16.7%; Fred. No. 0.00011;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXC 15  
DB 3858 CATTTTTATATC 3869

RESULT 49  
ABU96437  
ID ABU96437 standard; protein; 4440 AA.

XX ABU96437;

XX 25-JUL-2003 (first entry)

XX Novel human secreted and transmembrane protein #4.

XX Human; secreted and transmembrane protein; PRO; transgenic animal;

KW knockout; chromosome identification; tissue typing; tumour;

KW chondrocyte proliferation; chondrocyte differentiation;

XX tumor necrosis factor-alpha release stimulator.

OS Homo sapiens.

XX US2003036144-A1.  
XX 20-FEB-2003.  
XX

PF 01-JUL-2002; 2002US-00187601..  
XX 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 05-MAY-1998; 98US-0083559P.  
PR 06-MAY-1998; 98US-0084366P.  
PR 07-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0085709P.  
PR 22-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 02-JUN-1998; 98US-0087760P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.

PR 09-JUN-1998; 98US-0088655P.  
 PR 10-JUN-1998; 98US-0088722P.  
 PR 10-JUN-1998; 98US-0088738P.  
 PR 10-JUN-1998; 98US-0088740P.  
 PR 10-JUN-1998; 98US-0088811P.  
 PR 10-JUN-1998; 98US-0088824P.  
 PR 10-JUN-1998; 98US-0088825P.  
 PR 10-JUN-1998; 98US-0088826P.  
 PR 11-JUN-1998; 98US-0088861P.  
 PR 11-JUN-1998; 98US-0088863P.  
 PR 11-JUN-1998; 98US-0088876P.  
 PR 12-JUN-1998; 98US-0089090P.  
 PR 12-JUN-1998; 98US-0089105P.  
 PR 16-JUN-1998; 98US-0089512P.  
 PR 16-JUN-1998; 98US-0089514P.  
 PR 17-JUN-1998; 98US-0089538P.  
 PR 17-JUN-1998; 98US-0089588P.  
 PR 17-JUN-1998; 98US-0089653P.  
 PR 18-JUN-1998; 98US-0089908P.  
 PR 19-JUN-1998; 98US-0089952P.  
 PR 22-JUN-1998; 98US-0090246P.  
 PR 22-JUN-1998; 98US-0090252P.  
 PR 22-JUN-1998; 98US-0090254P.  
 PR 24-JUN-1998; 98US-0090429P.  
 PR 24-JUN-1998; 98US-0090435P.  
 PR 24-JUN-1998; 98US-0090444P.  
 PR 24-JUN-1998; 98US-0090461P.  
 PR 24-JUN-1998; 98US-0090535P.  
 PR 24-JUN-1998; 98US-0090540P.  
 PR 25-JUN-1998; 98US-0090676P.  
 PR 25-JUN-1998; 98US-0090688P.  
 PR 25-JUN-1998; 98US-0090689P.  
 PR 25-JUN-1998; 98US-0090690P.  
 PR 25-JUN-1998; 98US-0090694P.  
 PR 25-JUN-1998; 98US-0090695P.  
 PR 25-JUN-1998; 98US-0090696P.  
 PR 26-JUN-1998; 98US-00105413.  
 PR 26-JUN-1998; 98US-0090862P.  
 PR 26-JUN-1998; 98US-0090863P.  
 PR 26-JUN-1998; 98US-0091010P.  
 PR 01-JUL-1998; 98US-0091359P.  
 PR 01-JUL-1998; 98US-0091544P.  
 PR 02-JUL-1998; 98US-0091478P.  
 PR 02-JUL-1998; 98US-0091486P.  
 PR 02-JUL-1998; 98US-0091526P.  
 PR 02-JUL-1998; 98US-0091628P.  
 PR 02-JUL-1998; 98US-0091632P.  
 PR 04-JUL-1998; 98US-0094006P.  
 PR 04-AUG-1998; 98US-0095282P.  
 PR 10-AUG-1998; 98US-0095989P.  
 PR 10-AUG-1998; 98US-0096012P.  
 PR 17-AUG-1998; 98US-0096757P.  
 PR 17-AUG-1998; 98US-0096766P.  
 PR 17-AUG-1998; 98US-0096867P.  
 PR 17-AUG-1998; 98US-0096891P.  
 PR 17-AUG-1998; 98US-0096897P.  
 PR 18-AUG-1998; 98US-0096949P.  
 PR 18-AUG-1998; 98US-0096959P.  
 PR 18-AUG-1998; 98US-0097022P.  
 PR 26-AUG-1998; 98US-0097952P.  
 PR 26-AUG-1998; 98US-0097954P.  
 PR 26-AUG-1998; 98US-0097955P.  
 PR 26-AUG-1998; 98US-0097971P.  
 PR 26-AUG-1998; 98US-0097974P.  
 PR 26-AUG-1998; 98US-0098014P.  
 PR 01-SEP-1998; 98US-0098716P.  
 PR 01-SEP-1998; 98US-0098723P.  
 PR 02-SEP-1998; 98US-0098803P.  
 PR 02-SEP-1998; 98US-0098821P.  
 PR 02-SEP-1998; 98US-0098843P.  
 PR 09-SEP-1998; 98US-0099602P.  
 PR 10-SEP-1998; 98US-0099741P.  
 PR 10-SEP-1998; 98US-0099754P.

PR 10-SEP-1998; 98US-0099763P.  
 PR 10-SEP-1998; 98US-0099812P.  
 PR 15-SEP-1998; 98US-0100388P.  
 PR 16-SEP-1998; 98US-0100662P.  
 PR 16-SEP-1998; 98US-0100664P.  
 PR 16-SEP-1998; 98US-0101751P.  
 PR 16-SEP-1998; 98US-0101751P.  
 PR 17-SEP-1998; 98US-0100683P.  
 PR 17-SEP-1998; 98US-0100684P.  
 PR 17-SEP-1998; 98US-0100919P.  
 PR 17-SEP-1998; 98US-0100930P.  
 PR 18-SEP-1998; 98US-0100849P.  
 PR 18-SEP-1998; 98US-0101014P.  
 PR 18-SEP-1998; 98US-0101068P.  
 PR 23-SEP-1998; 98US-0101471P.  
 PR 23-SEP-1998; 98US-0101472P.  
 PR 23-SEP-1998; 98US-0101475P.  
 PR 23-SEP-1998; 98US-0101477P.  
 PR 24-SEP-1998; 98US-0101738P.  
 PR 24-SEP-1998; 98US-0101739P.  
 PR 24-SEP-1998; 98US-0101743P.  
 PR 24-SEP-1998; 98US-0101922P.  
 PR 25-SEP-1998; 98US-0101786P.  
 PR 29-SEP-1998; 98US-0102207P.  
 PR 29-SEP-1998; 98US-0102240P.  
 PR 29-SEP-1998; 98US-0102330P.  
 PR 29-SEP-1998; 98US-0102331P.  
 PR 30-SEP-1998; 98US-0102487P.  
 PR 30-SEP-1998; 98US-0102570P.  
 PR 30-SEP-1998; 98US-0102571P.  
 PR 01-OCT-1998; 98US-0102684P.  
 PR 01-OCT-1998; 98US-0102687P.  
 PR 02-OCT-1998; 98US-0102958P.  
 PR 06-OCT-1998; 98US-0103258P.  
 PR 06-OCT-1998; 98US-0103449P.

Query Match 52.9%; Score 18; DB 6; Length 4440;  
 Best Local Similarity 16.7%; Pred. No. 0.00011;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 15

Db 3858 CATTTTTAATAC 3869

RESULT 50

ID ABU99046 standard; protein; 4440 AA.

AC ABU99046;

DT 01-AUG-2003 (first entry)

DE Novel human secreted and transmembrane protein #4.

KW Human: secreted and transmembrane protein; PRO: cytosolic; gene therapy;  
 KW chondrocyte stimulator; tumour; adrenal tumour; lung tumour;  
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;  
 KW cervical tumour; liver tumour; TNF-alpha release;  
 KW tumour necrosis factor alpha release; chondrocyte cell proliferation;  
 KW chondrocyte cell differentiation; pharmaceutical; diagnostic; biosensor;  
 KW bioreactor.

OS Homo sapiens.

PN US2003013153-A1.

PD 16-JAN-2003.

PF 19-JUN-2002; 2002US-00175737.

XX 18-SEP-1997; 97US-0059263P.  
 XX 18-SEP-1997; 97US-0059266P.

PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065111P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083122P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084314P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088302P.  
PR 05-JUN-1998; 98US-0088312P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0088909P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 04-AUG-1998; 98US-0094006P.  
PR 10-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095988P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 26-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.

PR 16-SEP-1998; 98US-010664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98WC-US019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.

Query Match 52.9%; Score 18; DB 6; Length 4440;  
Best Local Similarity 16.7%; Pred. No. 0.00011;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXC 15  
Db 3858 CATTTTAATAC 3869

Search completed: January 4, 2006, 15:56:12  
Job time : 132.183 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 / Search time 18.6261 Seconds  
(without alignments)  
92.983 Million cell updates/sec

Title: US-09-932-322-5  
Perfect score: 34  
Sequence: 1 XXXXXXXXXXXXXXXXXX 18

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database :  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	52.9	307	1	GSFP3
2	17	50.0	245	2	T24565
3	16	47.1	152	2	D72630
4	16	47.1	244	2	T31838
5	16	47.1	248	2	T31841
6	16	47.1	257	2	T21029
7	16	47.1	539	2	T52565
8	16	47.1	667	2	A48579
9	16	47.1	770	2	T31024
10	16	47.1	1041	2	T31097
11	15	44.1	154	2	AC0416
12	15	44.1	226	2	B41378
13	15	44.1	248	2	T21786
14	15	44.1	268	2	A84746
15	15	44.1	273	2	S69193
16	15	44.1	283	2	AH2744
17	15	44.1	287	2	S57770
18	15	44.1	289	2	S49812
19	15	44.1	318	2	G97525
20	15	44.1	357	2	T21152
21	15	44.1	363	2	S18765
22	15	44.1	392	2	T27303
23	15	44.1	409	2	T08928
24	15	44.1	420	2	S74388
25	15	44.1	423	2	D96552
26	15	44.1	503	2	S31126
27	15	44.1	569	2	B87180
28	15	44.1	727	2	JC7818
29	15	44.1	779	1	S40382

30	15	44.1	1003	1	PVZAM	spheroidin precurs
31	15	44.1	1043	2	T19734	hypothetical prote
32	15	44.1	1186	2	T18210	delta endotoxin -
33	15	44.1	1364	2	T10236	xanthine dehydrog
34	15	44.1	1451	2	S65571	pattern formation
35	15	44.1	2098	2	T25888	hypothetical prote
36	15	44.1	2176	2	T13806	coucan gene protei
37	15	44.1	2187	2	S60224	polyketide synth
38	15	44.1	2533	2	T28675	alpha-51D immobili
39	15	44.1	2533	2	T28675	alpha-51D immobili
40	15	44.1	2543	2	T21687	surface antigen - P
41	15	44.1	2703	1	A24420	notch protein - fr
42	15	44.1	4767	2	T31345	hypothetical prote
43	15	44.1	7829	2	T15789	hypothetical prote
44	15	44.1	7962	2	I38346	elastic titin - hu
45	14	41.2	21	2	D22565	R-phycocerythrin be
46	14	41.2	88	2	G82754	hypothetical prote
47	14	41.2	92	2	I54781	fibroblast growth
48	14	41.2	97	2	T28947	hypothetical prote
49	14	41.2	100	2	C72775	hypothetical prote
50	14	41.2	101	2	T14780	hypothetical prote
51	14	41.2	103	2	S73440	hypothetical prote
52	14	41.2	107	2	S46078	probable membrane
53	14	41.2	107	4	S57388	hypothetical prote
54	14	41.2	108	2	S67638	probable membrane
55	14	41.2	129	2	S48641	lysosome (EC 3.2.1
56	14	41.2	130	2	S57896	polyketide synthas
57	14	41.2	131	2	A43980	neurophysin 2 [val
58	14	41.2	133	2	S44539	probable membrane
59	14	41.2	135	2	T15610	hypothetical prote
60	14	41.2	143	2	A54793	ponticulin precurs
61	14	41.2	154	2	B72513	hypothetical prote
62	14	41.2	156	2	T31839	hypothetical prote
63	14	41.2	156	2	AH0865	prepilin peptidase
64	14	41.2	156	2	C65065	prepilin peptidase
65	14	41.2	158	2	S35201	serine proteinase
66	14	41.2	159	2	S82692	hypothetical prote
67	14	41.2	164	2	T28741	hypothetical prote
68	14	41.2	165	2	T28742	hypothetical prote
69	14	41.2	166	2	T33970	hypothetical prote
70	14	41.2	177	2	S73288	phycocerythrin beta
71	14	41.2	177	2	S30932	R-phycocerythrin be
72	14	41.2	177	2	F72706	hypothetical prote
73	14	41.2	195	2	S39967	hypothetical prote
74	14	41.2	197	2	T28739	nitrate homology
75	14	41.2	199	2	S50363	hypothetical prote
76	14	41.2	204	2	T41853	AcMNPV orf115 - Bo
77	14	41.2	205	2	T10384	hypothetical prote
78	14	41.2	209	2	JC7239	peroxiredoxin V -
79	14	41.2	214	2	C70812	probable lipQ prot
80	14	41.2	225	2	T22417	hypothetical prote
81	14	41.2	232	2	A60095	larval glue protei
82	14	41.2	249	2	T31837	hypothetical prote
83	14	41.2	250	2	T31836	hypothetical prote
84	14	41.2	250	2	T31835	hypothetical prote
85	14	41.2	259	2	H71353	conserved hypotet
86	14	41.2	264	2	C70603	hypothetical prote
87	14	41.2	264	2	S42793	pds-1 protein - Ca
88	14	41.2	268	2	T05971	probable receptor -
89	14	41.2	270	2	T27786	hypothetical prote
90	14	41.2	274	2	D82784	phosphomethylprim
91	14	41.2	275	2	T05969	receptor-like kina
92	14	41.2	276	2	A80860	hypothetical prote
93	14	41.2	286	1	C42053	gap junction prote
94	14	41.2	288	2	T21158	hypothetical prote
95	14	41.2	293	2	S15524	variant surface gl
96	14	41.2	298	2	S15524	variant surface gl
97	14	41.2	300	2	T05729	hypothetical prote
98	14	41.2	300	2	S16344	variant surface gl
99	14	41.2	312	2	T31834	hypothetical prote
100	14	41.2	315	2	T34528	hypothetical prote
101	14	41.2	328	2	T16065	iron-sulfur cofact
102	14	41.2	329	2	T15823	hypothetical prote

103	14	41.2	337	2	147079	176	14	41.2	706	2	149700	related to Ap-1-1i
104	14	41.2	343	1	HMITVA	177	14	41.2	710	2	T21339	hypothetical prote
105	14	41.2	343	2	S45321	178	14	41.2	712	2	T18195	gag protein - silk
106	14	41.2	343	2	U02370	179	14	41.2	722	2	T00049	hemocyte protein A
107	14	41.2	343	2	U02372	180	14	41.2	729	2	S36605	nitrate reductase
108	14	41.2	343	2	U02371	181	14	41.2	742	2	T25415	hypothetical prote
109	14	41.2	344	1	A27701	182	14	41.2	758	2	A96687	hypothetical prote
110	14	41.2	344	1	A32141	183	14	41.2	769	2	A41029	integrin beta-8 ch
111	14	41.2	344	2	I45894	184	14	41.2	776	2	T02584	probable protein k
112	14	41.2	344	2	I57698	185	14	41.2	780	2	T03156	ribonucleoside-dip
113	14	41.2	344	2	P00319	186	14	41.2	791	2	B46171	coat protein - lei
114	14	41.2	344	2	P00321	187	14	41.2	794	2	B46171	hypothetical prote
115	14	41.2	345	2	A05279	188	14	41.2	856	2	D86258	protein F501.10 l
116	14	41.2	347	2	S12916	189	14	41.2	880	2	AE0179	probable ATPase ch
117	14	41.2	349	2	AD3650	190	14	41.2	882	2	AE0119	Cip ATPase limport
118	14	41.2	353	2	S16785	191	14	41.2	927	2	T21772	hypothetical prote
119	14	41.2	353	2	S65086	192	14	41.2	946	2	C86549	polymorphic outer
120	14	41.2	353	2	UC4875	193	14	41.2	946	2	D81594	polymorphic membra
121	14	41.2	354	2	A35788	194	14	41.2	946	2	C72075	polymorphic outer
122	14	41.2	357	2	S58168	195	14	41.2	988	2	T51054	related to alpha-a
123	14	41.2	358	2	S69886	196	14	41.2	989	2	T48845	insulin II gene en
124	14	41.2	364	2	S57937	197	14	41.2	1011	2	S75806	hemolysin secretio
125	14	41.2	370	2	S32173	198	14	41.2	1088	2	A69493	cysteine proteinas
126	14	41.2	378	2	D86307	199	14	41.2	1105	2	T18295	Ap-3 adaptor compl
127	14	41.2	384	2	D86598	200	14	41.2	1165	2	S58236	pyruvate (flavodox
128	14	41.2	384	2	D71478	201	14	41.2	1206	2	B86445	hypothetical prote
129	14	41.2	384	2	C81741	202	14	41.2	1229	2	S14199	probable adenylate
130	14	41.2	384	2	P72025	203	14	41.2	1230	2	B64664	outer membrane pro
131	14	41.2	385	2	T25452	204	14	41.2	1237	2	D71850	probable outer mem
132	14	41.2	390	2	B88925	205	14	41.2	1237	2	A34598	ecdysone-induced p
133	14	41.2	392	2	T46418	206	14	41.2	1268	2	G85154	hypothetical prote
134	14	41.2	424	2	S74788	207	14	41.2	1287	2	T42658	hypothetical prote
135	14	41.2	425	2	T48724	208	14	41.2	1313	2	T30548	adenylate cyclase
136	14	41.2	433	2	T50395	209	14	41.2	1331	2	T18310	receptor-adenylate
137	14	41.2	439	2	A36385	210	14	41.2	1332	2	T15670	hypothetical prote
138	14	41.2	469	1	UC1318	211	14	41.2	1344	2	AD2103	two-component hydr
139	14	41.2	465	2	C84783	212	14	41.2	1357	2	S61187	probable membrane
140	14	41.2	474	2	T27297	213	14	41.2	1361	2	C71403	hypothetical 15.1k
141	14	41.2	478	2	T24805	214	14	41.2	1367	1	S48478	glucan 1,4-alpha-g
142	14	41.2	478	2	T41408	215	14	41.2	1380	2	T18309	receptor-adenylate
143	14	41.2	479	2	S65466	216	14	41.2	1387	2	A96771	hypothetical prote
144	14	41.2	494	2	T51529	217	14	41.2	1394	2	B34598	ecdysone-induced p
145	14	41.2	495	1	S46284	218	14	41.2	1429	2	T13720	gene expanded prot
146	14	41.2	496	1	U02152	219	14	41.2	1443	2	S05979	steroid hormone re
147	14	41.2	499	2	S22571	220	14	41.2	1450	2	T30273	hypothetical prote
148	14	41.2	501	2	G85097	221	14	41.2	1797	2	T21889	hypothetical prote
149	14	41.2	515	2	A25048	222	14	41.2	1805	2	T21888	hypothetical prote
150	14	41.2	527	2	A98180	223	14	41.2	1820	2	S71853	genome polyprotein
151	14	41.2	527	2	AC3107	224	14	41.2	1823	2	S28974	vitellogenin precu
152	14	41.2	529	2	A44508	225	14	41.2	1986	2	S28353	probable polyketid
153	14	41.2	537	2	H88087	226	14	41.2	2109	2	T17490	polyketide synthas
154	14	41.2	559	1	A29941	227	14	41.2	2257	1	A57710	acetyl-CoA carboxy
155	14	41.2	559	1	A35029	228	14	41.2	2311	2	T06161	acetyl-CoA carboxy
156	14	41.2	561	2	UC2447	229	14	41.2	2325	2	T02235	acetyl-CoA carboxy
157	14	41.2	563	1	S49889	230	14	41.2	2395	1	S50820	surface protein ty
158	14	41.2	566	1	HMITV	231	14	41.2	2437	2	S42612	transmembrane prot
159	14	41.2	566	1	HMITVR	232	14	41.2	2458	2	T17420	probable polyketid
160	14	41.2	566	1	HMITVR	233	14	41.2	2531	2	T17420	hypothetical prote
161	14	41.2	566	2	S69887	234	14	41.2	2704	2	S09118	G surface protein
162	14	41.2	566	2	S69888	235	14	41.2	2718	2	A23475	G surface protein
163	14	41.2	567	1	B45865	236	14	41.2	3133	2	S52093	hemocytin - silkwo
164	14	41.2	575	1	VC1JHD	237	14	41.2	3164	1	WMBEH6	U36 protein - hum
165	14	41.2	593	2	T10301	238	14	41.2	3172	2	S22012	erythronolide synt
166	14	41.2	595	2	A42086	239	14	41.2	3178	2	S13595	6-deoxyerythronoli
167	14	41.2	598	2	T46327	240	14	41.2	3229	2	S27852	probable cell-surf
168	14	41.2	606	2	S35427	241	14	41.2	3345	2	T13423	hypothetical prote
169	14	41.2	606	2	D86232	242	14	41.2	3491	2	T43231	probable 6-deoxyer
170	14	41.2	620	2	T50232	243	14	41.2	4427	2	PN0637	polyketide synthas
171	14	41.2	649	2	C90113	244	13	38.2	39	2	B31940	microtubule-associ
172	14	41.2	653	2	S40962	245	13	38.2	55	1	MTNHB2	neurotoxin B-II -
173	14	41.2	661	2	D85361	246	13	38.2	57	2	B38180	hypothetical prote
174	14	41.2	686	2	A59348	247	13	38.2	61	1	T2NJ2Y	short toxin CM-2 -
175	14	41.2	697	1	TVPPTL	248	13	38.2	63	1	XTBO1	acrosin inhibitor



249	13	38.2	66	2	A28644	kappa-neurotoxin -	322	13	38.2	201	2	T16181	hypothetical prote
250	13	38.2	66	2	JN0408	neurotoxin Os-1 -	323	13	38.2	202	2	C90742	hypothetical prote
251	13	38.2	71	2	T00008	copy number contro	324	13	38.2	202	2	R85592	hypothetical prote
252	13	38.2	79	2	T27605	hypothetical prote	325	13	38.2	204	2	D72864	AcOrf-115 protein
253	13	38.2	81	2	B42465	probable transcrip	326	13	38.2	205	2	H71457	hypothetical prote
254	13	38.2	98	2	B84599	hypothetical prote	327	13	38.2	208	2	C97654	hypothetical prote
255	13	38.2	101	2	T42262	hypothetical prote	328	13	38.2	208	2	A12877	conserved hypotet
256	13	38.2	102	2	AR1357	hypothetical prote	329	13	38.2	209	2	A82680	2-amino-4-hydroxy-
257	13	38.2	102	2	AR1727	hypothetical prote	330	13	38.2	210	2	S75973	hypothetical prote
258	13	38.2	107	2	T27988	hypothetical prote	331	13	38.2	213	2	R75585	uracase accessory p
259	13	38.2	108	2	B95314	hypothetical prote	332	13	38.2	216	2	T50615	hypothetical prote
260	13	38.2	111	2	AC0733	probable exported	333	13	38.2	216	2	C90162	hypothetical prote
261	13	38.2	111	2	T20452	hypothetical prote	334	13	38.2	217	2	A86600	Ct734 hypothetical
262	13	38.2	111	2	T34565	hypothetical prote	335	13	38.2	217	2	A72024	conserved hypotet
263	13	38.2	112	2	T51142	nonspecific lipid-	336	13	38.2	217	2	A42644	neuB protein - Bac
264	13	38.2	112	2	T48770	hypothetical prote	337	13	38.2	217	2	T48642	hypothetical prote
265	13	38.2	114	2	S57692	probable membrane	338	13	38.2	217	2	T48785	hypothetical prote
266	13	38.2	114	2	F69264	hypothetical prote	339	13	38.2	220	2	T25789	L-thamose-binding
267	13	38.2	114	2	T01930	ethylene-response	340	13	38.2	222	2	T27939	hypothetical prote
268	13	38.2	117	2	T02169	homeobox protein H	341	13	38.2	225	2	S45356	probable serine pr
269	13	38.2	119	2	B45937	early chorion prot	342	13	38.2	226	2	G95287	probable [imported
270	13	38.2	119	2	S24292	chorion protein -	343	13	38.2	226	2	T21380	hypothetical prote
271	13	38.2	119	2	S24294	chorion protein -	344	13	38.2	227	2	A33937	Ig light chain (13
272	13	38.2	119	2	S24291	chorion protein -	345	13	38.2	227	2	A55472	ornithine decarbox
273	13	38.2	121	2	S24293	chorion protein -	346	13	38.2	229	1	MMVYPM	25k protein - pota
274	13	38.2	125	2	S30170	mercury resistance	347	13	38.2	233	2	T21496	hypothetical prote
275	13	38.2	126	2	AB0071	probable exported	348	13	38.2	235	2	T16916	hypothetical prote
276	13	38.2	127	2	S27915	nucleic acid-bindi	349	13	38.2	236	2	R84307	glucose 1-dehydrog
277	13	38.2	127	2	H90091	hypothetical prote	350	13	38.2	237	2	T25152	hypothetical prote
278	13	38.2	129	2	T47958	hypothetical prote	351	13	38.2	238	2	S77699	inner cell wall ma
279	13	38.2	132	1	MF1V2T	matrix protein M2	352	13	38.2	238	2	T22098	hypothetical prote
280	13	38.2	141	2	T33983	hypothetical prote	353	13	38.2	241	2	T24782	hypothetical prote
281	13	38.2	144	2	A36324	growth arrest-spec	354	13	38.2	242	2	A96794	unknown protein F1
282	13	38.2	145	2	H87436	hypothetical prote	355	13	38.2	242	2	H65136	hypothetical prote
283	13	38.2	145	2	T33037	hypothetical prote	356	13	38.2	243	2	F86006	hypothetical prote
284	13	38.2	146	1	W8BPT3	gene 19.2 protein	357	13	38.2	243	2	G91160	hypothetical prote
285	13	38.2	148	2	S19462	probable membrane	358	13	38.2	246	2	T00628	hypothetical prote
286	13	38.2	150	2	S66926	hypothetical prote	359	13	38.2	246	2	T25532	hypothetical prote
287	13	38.2	155	2	C62890	conserved hypotet	360	13	38.2	247	2	T21406	hypothetical prote
288	13	38.2	156	2	C91089	prepillin peptidase	361	13	38.2	247	2	S72859	hypothetical prote
289	13	38.2	156	2	D75258	hypothetical prote	362	13	38.2	248	2	S05007	calcium channel al
290	13	38.2	156	2	B85934	prepillin peptidase	363	13	38.2	248	2	T11913	hypothetical prote
291	13	38.2	157	2	JN0057	hypothetical 17K p	364	13	38.2	249	2	T21785	hypothetical prote
292	13	38.2	160	2	JN0503	peripherall myelin	365	13	38.2	249	2	T27782	hypothetical prote
293	13	38.2	160	2	S21721	growth arrest-spec	366	13	38.2	250	2	T36131	hypothetical prote
294	13	38.2	160	2	A41144	growth arrest-rele	367	13	38.2	250	2	T26725	hypothetical prote
295	13	38.2	162	2	T11978	phycocyanin alpha	368	13	38.2	252	2	T29527	hypothetical prote
296	13	38.2	163	2	D63883	hypothetical prote	369	13	38.2	254	2	E83619	probable transpor
297	13	38.2	163	2	B44813	probable molybdenu	370	13	38.2	255	2	T03408	MDS box protein -
298	13	38.2	166	1	NVBO2	vaopressin / neur	371	13	38.2	255	2	T34163	hypothetical prote
299	13	38.2	166	2	R81696	hypothetical prote	372	13	38.2	257	2	D71544	hypothetical prote
300	13	38.2	166	2	T32701	hypothetical prote	373	13	38.2	259	2	T27783	hypothetical prote
301	13	38.2	168	1	NVRT2	vaopressin / neur	374	13	38.2	261	2	S70006	finger protein zfo
302	13	38.2	168	2	B43755	vaopressin / neur	375	13	38.2	262	2	S01914	El membrane glycop
303	13	38.2	168	2	I38944	melanoma ubiquitou	376	13	38.2	262	2	G86938	conserved hypotet
304	13	38.2	171	2	H84923	hypothetical prote	377	13	38.2	265	2	T30059	hypothetical prote
305	13	38.2	173	2	F96543	hypothetical prote	378	13	38.2	266	2	T10816	throxine deiodina
306	13	38.2	173	2	A71144	hypothetical prote	379	13	38.2	268	2	T43486	hypothetical prote
307	13	38.2	176	2	T30961	hypothetical prote	380	13	38.2	269	2	T21407	hypothetical prote
308	13	38.2	180	2	A95283	lipuin peroxidase	381	13	38.2	271	2	E86352	protein T26F17.16
309	13	38.2	180	2	PS0011	hypothetical prote	382	13	38.2	272	2	T19418	hypothetical prote
310	13	38.2	180	2	E71337	hypothetical prote	383	13	38.2	274	2	T39087	hypothetical prote
311	13	38.2	182	2	S29611	pollen-preferentia	384	13	38.2	275	2	A34866	T-cell surface pro
312	13	38.2	183	2	T38923	microfibril-associ	385	13	38.2	275	2	S08464	T-cell allergen
313	13	38.2	183	2	A54151	microfibril-associ	386	13	38.2	275	2	T05822	hypothetical prote
314	13	38.2	183	2	A49313	microfibril-associ	387	13	38.2	276	2	T43541	purine nucleotide
315	13	38.2	185	2	S57419	PAR-1a protein - c	388	13	38.2	276	2	AD1987	hypothetical prote
316	13	38.2	185	2	S62698	photoassimilate-re	389	13	38.2	277	2	H97266	mind family ATPase
317	13	38.2	185	2	S62699	photoassimilate-re	390	13	38.2	278	2	T16494	hypothetical prote
318	13	38.2	198	2	JR0316	protoroan RNase, Dd	391	13	38.2	280	2	A42474	chitinase (BC 3.2,
319	13	38.2	199	2	T48099	hypothetical prote	392	13	38.2	287	1	A45168	probable 3-oxoacyl
320	13	38.2	199	2	T32267	hypothetical prote	393	13	38.2	287	2	T04236	xyloglucan endo-1,
321	13	38.2	201	2	S37847	hypothetical prote	394	13	38.2	289	2	AE2006	hypothetical prote

395	13	38.2	291	2	T28846	hypothetical prote	468	13	38.2	380	1	D65201	glycerol dehydroge
396	13	38.2	294	1	G65106	lipoprotein precu	469	13	38.2	380	2	B91238	glycerol dehydroge
397	13	38.2	294	2	AC0901	conserved hypotnet	470	13	38.2	380	2	F86085	glycerol dehydroge
398	13	38.2	294	2	D91134	lipoprotein precu	471	13	38.2	382	2	B86268	glut1 protein - mou
399	13	38.2	294	2	G85979	lipoprotein precu	472	13	38.2	384	2	S25771	gast1 protein - mou
400	13	38.2	294	2	T23916	hypothetical prote	473	13	38.2	389	2	AD0722	glutamate dehydrog
401	13	38.2	294	2	T23682	hypothetical prote	474	13	38.2	393	2	A56947	dual specificity p
402	13	38.2	296	2	JT0548	killer toxin KHR p	475	13	38.2	394	2	A56115	dual specificity p
403	13	38.2	297	2	S06267	surface antigen H	476	13	38.2	395	2	AB0814	probable alcohol p
404	13	38.2	300	2	T48147	hypothetical prote	477	13	38.2	396	2	B49475	Desert hedgehog pr
405	13	38.2	301	2	T05493	pathogenesis-relat	478	13	38.2	400	2	B48613	surface protein SU
406	13	38.2	301	2	T40593	cytoplasmic dynein	479	13	38.2	401	2	B87684	hypothetical prote
407	13	38.2	303	2	H71277	probable DNA adenl	480	13	38.2	403	2	A13256	hypothetical prote
408	13	38.2	303	2	T46405	hypothetical prote	481	13	38.2	404	2	S58830	probable membrane
409	13	38.2	303	2	B88990	protein C36C5.13 l	482	13	38.2	410	2	AB3546	antibucyraldehyde
410	13	38.2	309	2	S37934	hypothetical prote	483	13	38.2	411	2	A82760	cell division prot
411	13	38.2	311	2	H64483	5-methyltetrahydro	484	13	38.2	411	2	T17653	hypothetical prote
412	13	38.2	312	2	B75405	endopeptidase-rela	485	13	38.2	412	2	C38104	LFY floral meriste
413	13	38.2	313	2	AC0603	asparaginase (EC 3	486	13	38.2	415	2	C84829	hypothetical prote
414	13	38.2	314	1	A57126	dual specificity p	487	13	38.2	415	2	T43352	nuclear receptor N
415	13	38.2	314	2	B57126	dual specificity p	488	13	38.2	416	2	S65110	chitinase (EC 3.2.
416	13	38.2	314	2	A44437	regenerating liver	489	13	38.2	420	2	B38104	LFY floral meriste
417	13	38.2	317	2	T00500	probable elicitor	490	13	38.2	422	2	S35197	hypothetical prote
418	13	38.2	317	2	D86292	protein F7H2.14 l1	491	13	38.2	424	2	A38104	LFY floral meriste
419	13	38.2	321	2	D64820	probable asparagin	492	13	38.2	425	2	B84631	probable serine ca
420	13	38.2	323	2	S19650	cysteine proteinase	493	13	38.2	425	2	F70608	hypothetical prote
421	13	38.2	323	2	A39877	autocrine motility	494	13	38.2	426	2	F85574	probable symptom p
422	13	38.2	324	1	D42951	alkanal monooxygen	495	13	38.2	426	2	F90723	probable transport
423	13	38.2	324	2	S47432	cathepsin L (EC 3.	496	13	38.2	430	2	JC4720	toxin co-regulated
424	13	38.2	324	2	C86582	hypothetical prote	497	13	38.2	433	1	S24353	prolaseome 26S sub
425	13	38.2	324	2	D72041	conserved hypotnet	498	13	38.2	433	2	S53709	MSI1 protein homol
426	13	38.2	325	2	T27688	hypothetical prote	499	13	38.2	433	2	S39349	MSI1 protein homol
427	13	38.2	325	4	S61652	hypothetical prote	500	13	38.2	433	2	F86163	hypothetical prote
428	13	38.2	327	2	T41463	conserved hypotnet	501	13	38.2	434	2	S50865	avermectin-sensiti
429	13	38.2	329	2	S74227	cathepsin K (EC 3.	502	13	38.2	435	2	T20152	hypothetical prote
430	13	38.2	330	2	T25323	hypothetical prote	503	13	38.2	436	2	B81033	N-acetylglutamate
431	13	38.2	335	2	JE0115	zinc-finger protei	504	13	38.2	436	2	A81977	probable amino-ac
432	13	38.2	339	2	F47301	VirB11 homolog - B	505	13	38.2	437	2	S11497	branched-chain ami
433	13	38.2	341	1	T22930	phosphoprotein pho	506	13	38.2	437	2	H82295	branched chain ami
434	13	38.2	341	2	C70611	hemagglutinin prote	507	13	38.2	438	2	I38946	melanoma ubiquitou
435	13	38.2	344	2	QJ1643	hemagglutinin - in	508	13	38.2	440	2	D69596	branched-chain ami
436	13	38.2	346	1	MOH0D	N4-(beta-N-acetyl g	509	13	38.2	440	2	F97098	branched-chain ami
437	13	38.2	347	2	S12955	calcium channel pr	510	13	38.2	445	2	S37779	porin precursor, m
438	13	38.2	348	2	T14141	NADH2 dehydrogenas	511	13	38.2	445	2	C69596	branched-chain ami
439	13	38.2	349	2	B81855	adenine glycosylas	512	13	38.2	445	2	T18129	hypothetical prote
440	13	38.2	349	2	D81088	N/G-specific adeni	513	13	38.2	446	1	C70016	allantoinase homol
441	13	38.2	353	2	T27800	hypothetical prote	514	13	38.2	449	2	S50546	adenosylhomocyste
442	13	38.2	354	2	S70595	NADH2 dehydrogenas	515	13	38.2	453	2	T09745	myb-related protei
443	13	38.2	355	1	ADZM	fructose-bisphosph	516	13	38.2	454	2	D75446	cytochrome p450-11
444	13	38.2	355	2	T51414	CONSTANS-like 1 -	517	13	38.2	455	2	T48973	cytochrome p450-11
445	13	38.2	358	2	T47550	fructose bisphosph	518	13	38.2	455	2	AG0308	isochorismate synt
446	13	38.2	358	2	A84781	probable fructose	519	13	38.2	457	2	I48206	triacylglycerol 11
447	13	38.2	361	2	S59531	manganese peroxida	520	13	38.2	459	2	JC7931	cellulohydroxylase
448	13	38.2	362	2	T05824	probable pectin ac	521	13	38.2	461	2	S21223	triacylglycerol 11
449	13	38.2	363	2	JC5536	C-fringe protein 1	522	13	38.2	461	2	S50864	avermectin-sensiti
450	13	38.2	364	1	T18972	phosphoprotein pho	523	13	38.2	464	2	C46157	hexokinase (EC 2.7
451	13	38.2	364	2	C45574	40.7 kda polypepti	524	13	38.2	465	2	A46157	hexokinase (EC 2.7
452	13	38.2	364	2	H86384	40.1k. hypothetical	525	13	38.2	465	2	I49693	glucokinase (EC 2.
453	13	38.2	366	2	H89025	protein F13A2.6 l1	526	13	38.2	465	2	A31810	hexokinase (EC 2.7
454	13	38.2	367	1	S24411	dual specificity p	527	13	38.2	465	2	S41084	triacylglycerol 11
455	13	38.2	367	1	S29090	dual specificity p	528	13	38.2	466	2	B46157	hexokinase (EC 2.7
456	13	38.2	367	2	MOHU	alpha-2-HS-glycopr	529	13	38.2	467	2	B89605	protein F18G5.2 l1
457	13	38.2	367	1	AH0936	glycerol dehydroge	530	13	38.2	471	2	S54460	hypothetical prote
458	13	38.2	367	2	S52265	dual specificity p	531	13	38.2	472	1	A26730	ovoinhibitor precu
459	13	38.2	367	2	D82763	UDP-N-acetylglucos	532	13	38.2	474	2	B46746	glycine hydroxymet
460	13	38.2	367	2	JC6087	helix-loop-helix t	533	13	38.2	474	2	AF0759	codytrinic acid A,C
461	13	38.2	368	2	T07938	refin PFB1010w - m	534	13	38.2	475	2	A33696	glycine hydroxymet
462	13	38.2	371	2	H71600	probable acetyl-Co	535	13	38.2	476	2	A71302	conserved hypotnet
463	13	38.2	372	1	OPJGAP	lignin peroxidase	536	13	38.2	479	2	T49871	S-receptor kinase
464	13	38.2	372	2	B32322	lignin peroxidase	537	13	38.2	482	1	S71770	calcium-dependent
465	13	38.2	372	2	JH0156	brefeldin A estera	538	13	38.2	487	2	A43758	atomatic-L-amino-a
466	13	38.2	372	2	A53050	heat shock transcr	539	13	38.2	487	2	D70876	probable polyketid
467	13	38.2	377	2	T04213		540	13	38.2	488	2		

541	13	38.2	488	2	T30602	hypothetical prote	614	13	38.2	650	2	S44806	F10E9.6 protein -
542	13	38.2	489	2	H69059	dihydroxy-acid deh	615	13	38.2	654	2	C87791	protein B0207.12 (
543	13	38.2	494	2	C70940	probable cobo prot	616	13	38.2	656	2	D84359	thermosome subunit
544	13	38.2	495	2	T20754	hypothetical prote	617	13	38.2	656	2	E29826	hypothetical 72.4K
545	13	38.2	496	2	T33456	hypothetical prote	618	13	38.2	657	2	T52460	hypothetical prote
546	13	38.2	498	2	S12061	hexokinase (EC 2.7	619	13	38.2	659	2	A57532	Na+-dependent nucl
547	13	38.2	498	2	S27849	variant surface gl	620	13	38.2	663	2	T30621	hypothetical prote
548	13	38.2	504	2	T01603	hypothetical prote	621	13	38.2	664	2	S60062	hevin precursor -
549	13	38.2	504	2	T86343	T22111.12 protein	622	13	38.2	669	2	A46511	hypothetical prote
550	13	38.2	504	2	T16526	hypothetical prote	623	13	38.2	675	2	G84663	envelope protein -
551	13	38.2	506	2	S37583	RING finger protei	624	13	38.2	677	2	C42125	hypothetical prote
552	13	38.2	506	2	T29968	hypothetical prote	625	13	38.2	684	2	G84730	trophozoite cystei
553	13	38.2	509	2	A95985	probable sugar kin	626	13	38.2	690	2	T11749	Mutator-like trans
554	13	38.2	511	2	S58339	celluloglucan I sylv	627	13	38.2	691	2	T31552	transferrin - Acta
555	13	38.2	513	1	TVHURP	ret finger protein	628	13	38.2	700	2	S38426	hypothetical prote
556	13	38.2	517	2	T19962	hypothetical prote	629	13	38.2	713	2	A35502	chaperonin 60 prec
557	13	38.2	520	2	A96922	beta-glucosidase h	630	13	38.2	730	2	A75486	major surface-labe
558	13	38.2	530	2	JC7979	cellulohydrolase	631	13	38.2	733	1	A46373	hypothetical prote
559	13	38.2	536	2	T24000	hypothetical prote	632	13	38.2	738	2	D86345	probable serine/th
560	13	38.2	538	2	S52472	cell fusion protei	633	13	38.2	746	2	S52770	hypothetical prote
561	13	38.2	539	2	A29923	carboxylesterase (	634	13	38.2	757	2	JC7519	subtilisin-like pr
562	13	38.2	539	2	T46132	hypothetical prote	635	13	38.2	757	2	T16609	subtilisin-like se
563	13	38.2	540	2	A31584	carboxylesterase (	636	13	38.2	764	2	H83879	hypothetical prote
564	13	38.2	541	2	T33583	hypothetical prote	637	13	38.2	768	2	B41029	3-oxoacyl-(acyl)-ca
565	13	38.2	541	2	B82342	probable transcrip	638	13	38.2	773	1	JH0609	integrin beta-8 ch
566	13	38.2	543	2	T27000	hypothetical prote	639	13	38.2	775	2	T00962	protein-tyrosine-p
567	13	38.2	546	2	T49931	hypothetical prote	640	13	38.2	781	2	S43534	hypothetical prote
568	13	38.2	549	2	JX0054	carboxylesterase (	641	13	38.2	795	1	T34468	integrin beta3 - c
569	13	38.2	554	2	A39060	carboxylesterase (	642	13	38.2	796	1	JC1285	hypothetical prote
570	13	38.2	557	2	A47162	cholesterase B (E	643	13	38.2	799	2	A38308	protein-tyrosine-p
571	13	38.2	557	2	A48434	variant-specific s	644	13	38.2	802	1	A36065	integrin beta-5 ch
572	13	38.2	560	1	F69059	axialine-tRNA liga	645	13	38.2	807	2	C85025	protein-tyrosine-p
573	13	38.2	561	2	S71597	carboxylesterase (	646	13	38.2	814	2	T49207	hypothetical prote
574	13	38.2	561	2	S62788	carboxylesterase (	647	13	38.2	829	1	A47373	receptor kinase-1i
575	13	38.2	561	2	C64459	dihydroxy-acid deh	648	13	38.2	832	2	JC8051	protein-tyrosine-p
576	13	38.2	562	2	A55281	carboxylesterase (	649	13	38.2	840	2	S41218	protein tyrosine p
577	13	38.2	563	1	VCMMW7	env polypotein -	650	13	38.2	843	1	JQ2229	HRI protein - yea
578	13	38.2	564	2	S08199	carboxychrome-c3 hydr	651	13	38.2	848	1	T02053	DNA-directed DNA p
579	13	38.2	565	2	S10367	carboxylesterase (	652	13	38.2	852	2	T28790	S-receptor kinase
580	13	38.2	566	2	S19307	carboxylesterase (	653	13	38.2	852	2	T28790	guanine nucleotide
581	13	38.2	566	2	S69889	hemagglutinin prec	654	13	38.2	885	2	JC7898	hypothetical prote
582	13	38.2	567	1	A41010	carboxylesterase (	655	13	38.2	891	2	T21942	3',5'-cyclic nucle
583	13	38.2	567	2	T08405	hypothetical prote	656	13	38.2	892	2	A41697	hypothetical prote
584	13	38.2	569	2	S50771	hexose transport p	657	13	38.2	895	2	T23191	nitrate assimilat
585	13	38.2	571	1	S30253	GABA transport pro	658	13	38.2	907	2	T27317	hypothetical prote
586	13	38.2	573	2	T27671	hypothetical prote	659	13	38.2	917	1	S15885	hexokinase (EC 2.7
587	13	38.2	579	2	T30635	hypothetical prote	660	13	38.2	917	2	JC2025	hexokinase (EC 2.7
588	13	38.2	581	2	T24393	hypothetical prote	661	13	38.2	931	2	T25148	hypothetical prote
589	13	38.2	582	1	VCVDAR	env polypotein -	662	13	38.2	934	2	G86548	polymorphic outer
590	13	38.2	591	2	I48141	acroganin - guine	663	13	38.2	934	2	A72075	polymorphic outer
591	13	38.2	594	2	H48613	env polypotein pr	664	13	38.2	950	2	F86222	hypothetical prote
592	13	38.2	596	2	T30498	probable ribonucle	665	13	38.2	951	2	T00260	hypothetical prote
593	13	38.2	601	2	D89711	protein F40E10.4 (	666	13	38.2	952	2	D81593	polymorphic membra
594	13	38.2	601	2	T22025	hypothetical prote	667	13	38.2	962	2	JC5571	subtilisin-like pr
595	13	38.2	602	1	S25316	hydroxymethylgluta	668	13	38.2	966	2	E84053	penicillin-binding
596	13	38.2	602	2	T37254	acetylcholinestera	669	13	38.2	969	1	A39490	subtilisin-like pr
597	13	38.2	605	2	S48940	hypothetical prote	670	13	38.2	975	2	JC5570	subtilisin-like pr
598	13	38.2	609	2	JC7819	metalloproteinae	671	13	38.2	983	2	T39902	translational elonga
599	13	38.2	616	2	T29234	hypothetical prote	672	13	38.2	993	2	S35633	DNA-binding protei
600	13	38.2	618	2	G82340	phosphoglucanate s	673	13	38.2	994	2	H96510	probable disease r
601	13	38.2	625	2	T10661	serine/threonine-d	674	13	38.2	1011	2	T50344	poly(A)+ RNA trans
602	13	38.2	626	2	B97358	glucose-inhibited	675	13	38.2	1012	2	T52000	poly(A)+ RNA trans
603	13	38.2	632	2	S58152	hypothetical prote	676	13	38.2	1016	2	T00375	hypothetical prote
604	13	38.2	632	2	T21602	hypothetical prote	677	13	38.2	1046	2	A26838	prestalk protein p
605	13	38.2	633	1	T24898	matrix glycoprotei	678	13	38.2	1060	2	T30823	bumetanide sensiti
606	13	38.2	634	1	GERTX1	hypothetical prote	679	13	38.2	1073	2	E83051	carbamoylphosphate
607	13	38.2	635	2	S76371	hypothetical prote	680	13	38.2	1074	2	JC5928	semaphorin F precu
608	13	38.2	636	2	H87789	protein C34G6.1 (i	681	13	38.2	1082	2	T31112	Aspase 2 (EC 3.6.1
609	13	38.2	638	2	T26490	hypothetical prote	682	13	38.2	1087	2	S02035	period clock prote
610	13	38.2	640	1	HHKW7A	dnak-type molecula	683	13	38.2	1093	2	T50652	Ap-3 complex betas
611	13	38.2	640	1	T07923	acetyl-CoA carboxy	684	13	38.2	1094	2	T50651	Ap3-complex beta-3
612	13	38.2	640	2	AB2251	glucose inhibited	685	13	38.2	1101	2	T16840	hypothetical prote
613	13	38.2	640	2	S62747	homeotic protein A	686	13	38.2	1105	2	C87622	TonB-dependent rec

687	13	38.2	1112	2	H97050	DNA Polymerase III	760	13	38.2	2143	2	JH0427	voltage-dependent
688	13	38.2	1129	2	T42732	A-kinase anchoring	761	13	38.2	2143	2	G96595	hypothetical prote
689	13	38.2	1149	2	I38006	M130 antigen precu	762	13	38.2	2166	2	G11339	calcium channel pr
690	13	38.2	1151	2	I38004	M130 antigen precu	763	13	38.2	2168	2	T30171	calcium channel pr
691	13	38.2	1156	2	I38005	M130 antigen precu	764	13	38.2	2171	2	S05054	calcium channel al
692	13	38.2	1159	2	I38465	probable potassium	765	13	38.2	2180	2	T29764	hypothetical prote
693	13	38.2	1160	2	T13713	betalactam protein - fr	766	13	38.2	2188	2	A70984	probable polyketid
694	13	38.2	1162	2	T21557	hypothetical prote	767	13	38.2	2195	2	T34264	hypothetical prote
695	13	38.2	1166	2	T13958	synGAP-bl protein	768	13	38.2	2201	2	S73014	polyketide synthas
696	13	38.2	1170	2	I45914	integrin alpha 2 s	769	13	38.2	2204	1	RKNZNV	genome polypeptide
697	13	38.2	1176	2	T94482	hypothetical prote	770	13	38.2	2220	2	A45290	calcium channel pr
698	13	38.2	1180	2	A35854	integrin alpha-1 c	771	13	38.2	2233	2	T28669	surface protein 51
699	13	38.2	1181	2	A33998	hypothetical prote	772	13	38.2	2240	2	T37057	probable multi-dom
700	13	38.2	1205	2	T41987	probable adenylyate	773	13	38.2	2257	2	D86483	protein F5J5.19 [i
701	13	38.2	1242	2	S14201	Rae-GTPase activat	774	13	38.2	2257	2	T09538	acetyl-CoA carboxy
702	13	38.2	1249	2	T14270	Ran-binding protei	775	13	38.2	2261	2	T07084	probable acetyl-Co
703	13	38.2	1265	2	S57968	transferrin-like p	776	13	38.2	2304	2	T07920	polyketide synthas
704	13	38.2	1271	2	T24008	cysteine rich prot	777	13	38.2	2359	2	E86483	notch-1 protein -
705	13	38.2	1274	2	T10729	ras GTPase-activa	778	13	38.2	2518	2	A12140	notch-1 protein -
706	13	38.2	1293	2	T42017	furin (EC 3.4.21.7	779	13	38.2	2524	2	A35844	polyketide synthas
707	13	38.2	1299	2	T43251	connectin 3B - chl	780	13	38.2	2531	2	S18188	variant-specific s
708	13	38.2	1323	2	PN0568	protein let-23 [im	781	13	38.2	2531	2	A46019	hypothetical prote
709	13	38.2	1323	2	E88257	RNA2 polypeptide -	782	13	38.2	2543	2	F69679	fibritillin-2 precu
710	13	38.2	1323	2	S06187	calcium channel al	783	13	38.2	2555	2	A40043	hypothetical prote
711	13	38.2	1331	2	S05011	protein F21D18.16	784	13	38.2	2610	2	T20968	protein HMWP1 - ye
712	13	38.2	1334	2	D96521	xanthine dehydroge	785	13	38.2	2706	2	T28155	lysine-specific s
713	13	38.2	1340	2	T10235	alpha-glycan synth	786	13	38.2	2844	2	S28291	hypothetical prote
714	13	38.2	1359	2	T43433	protein-cytosine x	787	13	38.2	2918	2	A54105	fibritillin-2 precu
715	13	38.2	1369	2	S70712	hypothetical prote	788	13	38.2	2946	2	T15840	hypothetical prote
716	13	38.2	1374	2	T20406	probable polyketid	789	13	38.2	3161	2	T30342	protein HMWP1 - ye
717	13	38.2	1391	2	D70634	probable membrane	790	13	38.2	3163	2	AB0223	probable polyketid
718	13	38.2	1402	2	S59823	protein ZC84.1 [im	791	13	38.2	3163	2	T17440	Ran-binding protei
719	13	38.2	1407	2	E88550	polyketide synthas	792	13	38.2	3224	1	S58870	reelin precursor -
720	13	38.2	1416	2	S73013	hypothetical prote	793	13	38.2	3461	2	T34513	trithorax protein
721	13	38.2	1446	2	T26327	breast cancer tumo	794	13	38.2	3507	2	T13857	mycosubactin synth
722	13	38.2	1461	2	T42632	polyketide synthas	795	13	38.2	3828	2	T44806	connectin/titin -
723	13	38.2	1484	2	AG2136	probable ppas prot	796	13	38.2	3971	2	T42215	zonadhesin - mouse
724	13	38.2	1488	2	C70984	polyketide synthas	797	13	38.2	4162	2	T42215	polyketide synthas
725	13	38.2	1489	2	S73015	probable ppas prot	798	13	38.2	5376	2	T30283	CAP5 protein - ant
726	13	38.2	1538	2	E70874	polyketide synthas	799	13	38.2	6420	2	S55030	neurotoxin Bmk 41-
727	13	38.2	1540	2	H87203	polyketide synthas	800	12	35.3	27	2	A59356	metallothionein Mt
728	13	38.2	1565	2	AD2135	antigen containing	801	12	35.3	35	2	S59356	hypothetical prote
729	13	38.2	1567	2	T03730	hypothetical prote	802	12	35.3	40	1	SMF	leech-derived tryp
730	13	38.2	1570	2	AC2012	hypothetical prote	803	12	35.3	40	1	B61194	conotoxin-like pro
731	13	38.2	1587	2	AB2012	sol protein, large	804	12	35.3	43	2	B44010	ovomucoid, third d
732	13	38.2	1597	1	BF9FSL	gene small optic 1	805	12	35.3	45	1	XKPO2A	ovomucoid, third d
733	13	38.2	1597	2	T08428	protein F11C7.4 [i	806	12	35.3	46	2	S50015	hypothetical prote
734	13	38.2	1722	2	E89753	hypothetical synthas	807	12	35.3	50	2	F82539	ovomucoid, third d
735	13	38.2	1743	2	T26859	hypothetical prote	808	12	35.3	52	2	T10299	ovomucoid, third d
736	13	38.2	1744	2	T17421	hypothetical synthas	809	12	35.3	54	2	A31436	ovomucoid, third d
737	13	38.2	1795	2	D97312	polyketide synthas	810	12	35.3	54	2	I61589	ovomucoid, third d
738	13	38.2	1815	2	S73021	polyketide synthas	811	12	35.3	54	2	I61588	ovomucoid, third d
739	13	38.2	1822	2	F87203	probable polyketid	812	12	35.3	54	2	F31439	ovomucoid, third d
740	13	38.2	1827	1	RRMP7M	genome polypeptide	813	12	35.3	54	2	T31446	ovomucoid, third d
741	13	38.2	1844	2	S01956	hypothetical prote	814	12	35.3	54	2	D61589	ovomucoid, third d
742	13	38.2	1871	2	A87204	probable ppas prot	815	12	35.3	54	2	B61492	ovomucoid, third d
743	13	38.2	1876	2	C70749	agrin - rat	816	12	35.3	54	2	G31442	ovomucoid, third d
744	13	38.2	1876	2	C70749	agrin - rat	817	12	35.3	54	2	G31442	ovomucoid, third d
745	13	38.2	1876	2	C70749	agrin - rat	818	12	35.3	54	2	G31442	ovomucoid, third d
746	13	38.2	1876	2	C70749	agrin - rat	819	12	35.3	54	2	G31442	ovomucoid, third d
747	13	38.2	1876	2	C70749	agrin - rat	820	12	35.3	54	2	G31442	ovomucoid, third d
748	13	38.2	1876	2	C70749	agrin - rat	821	12	35.3	54	2	G31442	ovomucoid, third d
749	13	38.2	1876	2	C70749	agrin - rat	822	12	35.3	54	2	G31442	ovomucoid, third d
750	13	38.2	1876	2	C70749	agrin - rat	823	12	35.3	54	2	G31442	ovomucoid, third d
751	13	38.2	1876	2	C70749	agrin - rat	824	12	35.3	54	2	G31442	ovomucoid, third d
752	13	38.2	1876	2	C70749	agrin - rat	825	12	35.3	54	2	G31442	ovomucoid, third d
753	13	38.2	1876	2	C70749	agrin - rat	826	12	35.3	54	2	G31442	ovomucoid, third d
754	13	38.2	1876	2	C70749	agrin - rat	827	12	35.3	54	2	G31442	ovomucoid, third d
755	13	38.2	1876	2	C70749	agrin - rat	828	12	35.3	54	2	G31442	ovomucoid, third d
756	13	38.2	1876	2	C70749	agrin - rat	829	12	35.3	54	2	G31442	ovomucoid, third d
757	13	38.2	1876	2	C70749	agrin - rat	830	12	35.3	54	2	G31442	ovomucoid, third d
758	13	38.2	1876	2	C70749	agrin - rat	831	12	35.3	54	2	G31442	ovomucoid, third d
759	13	38.2	1876	2	C70749	agrin - rat	832	12	35.3	54	2	G31442	ovomucoid, third d

833	12	35.3	54	2	B31443	ovomucoid, third d	906	12	35.3	111	2	B85888	detox protein (imp
834	12	35.3	54	2	A31439	ovomucoid, third d	907	12	35.3	112	2	JC7528	guanylate cyclase
835	12	35.3	54	2	G61492	ovomucoid (PSTI-ty	908	12	35.3	111	1	XLHU	collipase precursor
836	12	35.3	54	2	B31436	ovomucoid, third d	909	12	35.3	112	2	A46717	collipase precursor
837	12	35.3	54	2	B31436	ovomucoid, third d	910	12	35.3	112	2	S64448	hypothetical prote
838	12	35.3	55	2	B31442	ovomucoid, third d	911	12	35.3	113	2	F95248	conserved hypotnet
839	12	35.3	56	2	D31445	ovomucoid, third d	912	12	35.3	113	2	C98113	conserved hypotnet
840	12	35.3	56	2	A31446	ovomucoid, third d	913	12	35.3	115	2	S66924	probable membrane
841	12	35.3	56	2	C61588	ovomucoid (PSTI-ty	914	12	35.3	115	2	G95059	hypothetical prote
842	12	35.3	56	2	G98233	hypothetical prote	915	12	35.3	115	2	B96654	probable RING zinc
843	12	35.3	56	2	A25356	bdellin B-3 - medi	916	12	35.3	115	2	S45920	hypothetical prote
844	12	35.3	59	2	A47218	carboxyl ester lip	917	12	35.3	119	1	FSN03M	phospholipase A2 (
845	12	35.3	60	2	I61886	zinc finger protei	918	12	35.3	121	1	QOEC13	hypothetical 13.5K
846	12	35.3	60	2	B86123	hypothetical prote	919	12	35.3	121	2	PH1661	Ig heavy chain V r
847	12	35.3	62	2	S62862	toxin III-8 - Tily	920	12	35.3	121	2	T02926	acyl carrier prote
848	12	35.3	62	2	S62866	toxin III-8 - Tily	921	12	35.3	121	2	A98089	hypothetical prote
849	12	35.3	62	2	B39510	toxin III-8 - Braz	922	12	35.3	121	2	C85934	hypothetical prote
850	12	35.3	62	2	S54336	metallothionein-2a	923	12	35.3	121	2	Q01058	DNA-binding protei
851	12	35.3	62	2	S54335	metallothionein-2c	924	12	35.3	122	2	T18564	hypothetical prote
852	12	35.3	64	2	S66481	hypothetical prote	925	12	35.3	123	2	S30530	Ig heavy chain V r
853	12	35.3	66	2	E70083	hypothetical prote	926	12	35.3	124	2	T01251	hypothetical prote
854	12	35.3	67	2	JN0378	neurotoxin Os-3 -	927	12	35.3	125	1	NFB01	oxytocin / neuroph
855	12	35.3	67	2	S73050	probable acetyltra	928	12	35.3	125	1	NFHU1	oxytocin / neuroph
856	12	35.3	72	1	N2EPIJ	long neurotoxin 1	929	12	35.3	125	1	NFPG1	oxytocin / neuroph
857	12	35.3	72	1	F82825	hypothetical prote	930	12	35.3	125	1	NFR11	oxytocin / neuroph
858	12	35.3	73	1	N2OH2	long neurotoxin 2	931	12	35.3	125	1	NFSH1	oxytocin / neuroph
859	12	35.3	77	2	B82617	hypothetical prote	932	12	35.3	125	2	A43755	oxytocin / neuroph
860	12	35.3	78	2	E70891	hypothetical prote	933	12	35.3	128	2	S26786	Ig heavy chain V r
861	12	35.3	82	2	S70807	hypothetical prote	934	12	35.3	128	2	U02180	hypothetical 14.8K
862	12	35.3	84	1	UJ0152	acrosin/trypsin in	935	12	35.3	128	2	Q01739	hypothetical 14.7K
863	12	35.3	85	2	D34770	ORP4 protein - sai	936	12	35.3	128	2	S27917	hypothetical prote
864	12	35.3	86	2	I61885	zinc finger protei	937	12	35.3	129	2	S01661	lysozyme (EC 3.2.1
865	12	35.3	89	2	DA3692	T33 protein - rabb	938	12	35.3	129	2	S07435	lysozyme (EC 3.2.1
866	12	35.3	89	2	S71555	proteinase inhibit	939	12	35.3	130	2	G71212	hypothetical prote
867	12	35.3	89	2	UQ2361	wheat aluminum ind	940	12	35.3	131	2	I40656	hypothetical prote
868	12	35.3	92	2	T10783	protein mcl - Mont	941	12	35.3	132	2	F64487	hypothetical prote
869	12	35.3	93	2	T06470	probable chitinase	942	12	35.3	132	2	T15955	hypothetical prote
870	12	35.3	94	2	T03285	colipase B precurs	943	12	35.3	133	2	T37311	hypothetical prote
871	12	35.3	96	1	XLHOB	keratin high-sulfu	944	12	35.3	133	2	JH0270	lysozyme (EC 3.2.1
872	12	35.3	97	1	KRGTHM	keratin high-sulfu	945	12	35.3	133	2	JB0311	serine proteinase
873	12	35.3	97	1	KRSHH2	keratin high-sulfu	946	12	35.3	135	2	AF0865	conserved hypotnet
874	12	35.3	97	2	T42335	hypothetical prote	947	12	35.3	135	2	S36197	hypothetical prote
875	12	35.3	97	2	JU0041	hypothetical prote	948	12	35.3	135	2	G96691	hypothetical prote
876	12	35.3	97	2	D82789	hypothetical prote	949	12	35.3	136	2	T36241	hypothetical prote
877	12	35.3	97	2	C95984	hypothetical prote	950	12	35.3	137	1	JC4233	lysozyme (EC 3.2.1
878	12	35.3	98	1	S25283	amebapore A precu	951	12	35.3	137	2	T37213	hypothetical prote
879	12	35.3	98	1	KRSHH3	keratin high-sulfu	952	12	35.3	137	2	G84174	hypothetical prote
880	12	35.3	98	1	KRSHH4	keratin high-sulfu	953	12	35.3	138	2	AD0281	conserved hypotnet
881	12	35.3	98	2	S26913	Ig heavy chain V r	954	12	35.3	138	2	S34725	regulation protein
882	12	35.3	98	2	I47083	BLIIB2 high-sulfu	955	12	35.3	139	2	T29955	lysozyme (EC 3.2.1
883	12	35.3	98	2	I47086	BLIIB2 high-sulfu	956	12	35.3	140	2	JC5003	lysozyme (EC 3.2.1
884	12	35.3	98	2	B86719	hypothetical prote	957	12	35.3	140	2	T49347	hypothetical prote
885	12	35.3	100	2	F95897	hypothetical prote	958	12	35.3	140	2	T05525	hypothetical prote
886	12	35.3	101	2	F75500	hypothetical prote	959	12	35.3	141	2	T33710	hypothetical prote
887	12	35.3	101	2	T26641	hypothetical prote	960	12	35.3	141	2	PC1294	trophozoite surfac
888	12	35.3	101	2	F72544	hypothetical prote	961	12	35.3	142	2	AC1934	hypothetical prote
889	12	35.3	102	2	PH1249	Ig heavy chain V r	962	12	35.3	144	2	A89472	protein B0294.1 [i
890	12	35.3	102	2	B89864	conserved hypotnet	963	12	35.3	145	2	S74282	hypothetical prote
891	12	35.3	102	2	S50530	hypothetical prote	964	12	35.3	145	2	S02027	hemoglobin beta ch
892	12	35.3	102	2	A72604	hypothetical prote	965	12	35.3	145	2	G96581	F511.18 [imported
893	12	35.3	104	2	I50808	neuropeptide Y pre	966	12	35.3	145	2	T34539	hypothetical prote
894	12	35.3	104	2	T43968	hypothetical prote	967	12	35.3	146	2	B82568	conserved hypotnet
895	12	35.3	104	2	F45681	orf 61.2 - phage T	968	12	35.3	146	2	B95363	protein [imported
896	12	35.3	106	2	F72043	conserved hypotnet	969	12	35.3	147	2	S34148	thyrotropin beta c
897	12	35.3	106	2	C86580	Ct57 hypothetical	970	12	35.3	147	2	T47192	Ig heavy chain var
898	12	35.3	107	2	T51038	hypothetical prote	971	12	35.3	148	2	H71644	hypothetical prote
899	12	35.3	107	2	A82642	conserved plasmid	972	12	35.3	148	2	B49727	DNA-binding protei
900	12	35.3	108	2	E70931	hypothetical prote	973	12	35.3	148	2	T21334	hypothetical prote
901	12	35.3	108	2	AD0715	conserved hypotnet	974	12	35.3	149	2	JC5396	helix-loop-helix p
902	12	35.3	108	2	AD0715	asparaginopepsin I	975	12	35.3	151	2	T36314	hypothetical prote
903	12	35.3	109	2	SS4446	prepilin-like prot	976	12	35.3	152	1	KRSHHC	keratin high-sulfu
904	12	35.3	111	2	G91043	detox protein (imp	977	12	35.3	152	2	S57427	cysteine proteinas
905	12	35.3	111	2	H65020	ccha protein - Bsc	978	12	35.3	152	2	I47111	high-sulfur wool m

979 12 35.3 152 2 147109 high-sulfur wool m  
980 12 35.3 152 2 147112 high-sulfur wool m  
981 12 35.3 152 2 C39364 finger protein HTP  
982 12 35.3 153 2 C70958 hypothetical prote  
983 12 35.3 153 2 A96751 hypothetical prote  
984 12 35.3 154 2 S36584 B6 protein - human  
985 12 35.3 154 2 A49727 DNA-binding protei  
986 12 35.3 154 2 S47524 gene Id1 protein -  
987 12 35.3 155 1 T1ZM1 trypsin/factor XII  
988 12 35.3 155 2 JCS395 helix-loop-helix p  
989 12 35.3 155 2 D72761 hypothetical prote  
990 12 35.3 155 2 UC7732 trypsin-plasmin in  
991 12 35.3 157 2 S09805 hypothetical prote  
992 12 35.3 158 2 B34132 vasotocin / neurop  
993 12 35.3 158 2 G86964 hypothetical prote  
994 12 35.3 158 2 T26444 hypothetical prote  
995 12 35.3 159 2 D4615 coagulation factor  
996 12 35.3 159 2 UC0136 hypothetical 17.8k  
997 12 35.3 159 2 JQ1777 Salp2R protein pre  
998 12 35.3 159 2 T3696 hypothetical prote  
999 12 35.3 160 1 ERADA5 early E3 18.5K gly  
1000 12 35.3 160 2 D48232 cysteine-rich exte

## ALIGNMENTS

## RESULT 1

GSFP3 salivary glue protein *sgs-3* - fruit fly (*Drosophila melanogaster*)

C/Species: *Drosophila melanogaster*

C/Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 09-Jul-2004

C/Accession: A03329

R.Garfinkel, M.D.; Pruitt, R.E.; Meyerowitz, E.M.

J. Mol. Biol. 168, 765-789, 1983

A/Title: DNA sequences, gene regulation and modular protein evolution in the *Drosophila*

A/Reference number: A92904; MUID:83294545; PMID:6411930

A/Accession: A03329

A/Molecule type: DNA

A/Residues: 1-307 <GAR>

A/Cross-references: UNIPROT:P02840; UNIPARC:UPI00001245F4; GB:X01918; NID:98581; PIDN:CA

C/Comment: This protein is produced by third-instar larvae.

C/Genetics:

A/Gene: *sgs-3*

A/Cross-references: FlyBase:FBgn000373

A/Map position: 3L (68C)

A/Intons: 10/1

C/Superfamily: salivary glue protein

C/Keywords: salivary gland; tandem repeat

Query Match 52.9%; Score 18; DB 1; Length 307;  
Best Local Similarity 16.7%; Pred. No. 1.3e-10;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 44 CTTTTTTTTTTC 55

## RESULT 2

T24565 hypothetical protein T06C12.14 - *Caenorhabditis elegans*

C/Species: *Caenorhabditis elegans*

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T24565

R/Kelly, P.

Submitted to the EMBL Data Library, October 1996

A/Reference number: Z19908

A/Accession: T24565

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-245 <WIL>

A/Cross-references: UNIPROT:O18039; UNIPARC:UPI000007C269; EMBL:Z81116; PIDN:CAB03300.1;

A/Experimental source: clone T06C12  
C/Genetics:  
A/Gene: CESP:T06C12.14  
A/Map position: 5  
A/Intons: 33/3; 121/1  
C/Superfamily: *Caenorhabditis elegans* hypothetical protein C49G7.3

Query Match 50.0%; Score 17; DB 2; Length 245;  
Best Local Similarity 16.7%; Pred. No. 5.8e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 193 CASATTTTSSSTC 204

## RESULT 3

D72630 hypothetical protein APE1499 - *Aeropyrum pernix* (strain K1)

C/Species: *Aeropyrum pernix*

C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C/Accession: D72630

R.Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahara, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyr*

A/Reference number: A72450; MUID:99310339; PMID:10382966

A/Accession: D72630

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-152 <KAM>

A/Cross-references: UNIPROT:Q9YBV0; UNIPARC:UPI000005DF8F; DBJ:AF000061; NID:95104821; I

A/Experimental source: strain K1

A/Gene: APE1499

Query Match 47.1%; Score 16; DB 2; Length 152;  
Best Local Similarity 16.7%; Pred. No. 2.5e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 137 CSSTATTTLSTRC 148

## RESULT 4

T31838 hypothetical protein T05B4.12 - *Caenorhabditis elegans*

C/Species: *Caenorhabditis elegans*

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T31838

R.Bradsheaw, H.

Submitted to the EMBL Data Library, July 1997

A/Description: The sequence of *C. elegans* cosmid T05B4.

A/Reference number: Z21092

A/Accession: T31838

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-244 <BRA>

A/Cross-references: UNIPROT:O16421; UNIPARC:UPI000007AC7E; EMBL:AF016445; PIDN:AA069052.1

A/Experimental source: strain Bristol N2; clone T05B4

A/Gene: CESP:T05B4.12

C/Accession: T31838

A/Intons: 30/3; 116/1

C/Superfamily: *Caenorhabditis elegans* hypothetical protein C49G7.3

Query Match 47.1%; Score 16; DB 2; Length 244;  
Best Local Similarity 16.7%; Pred. No. 2.8e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 203 CTSYADSTSC 214

RESULT 5

T31841 hypothetical protein T05B4.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T31841

R/Bradshaw, H.

submitted to the EMBL Data Library, July 1997

A/Description: The sequence of C. elegans cosmid T05B4.

A/Reference number: Z21092

A/Accession: T31841

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-248 <BRA>

A/Cross-references: UNIPROT:O1424; UNIPARC:UPI00007P78C; EMBL:AF016445; PIDN:AAC69054.

A/Experimental source: strain Bristol N2; clone T05B4

C/Genetics:

A/Gene: CESP:T05B4.3

A/Map position: 5

A/Introns: 30/3; 118/1

C/Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3

Query Match

Best Local Similarity 47.1%; Score 16; DB 2; Length 248;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 207 CTSFADSSSC 218

RESULT 6

T21029 hypothetical protein F16H6.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T21029

R/Mathews, L.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z19362

A/Accession: T21029

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-257 <WTL>

A/Cross-references: UNIPROT:Q9XV90; UNIPARC:UPI000075F8F; EMBL:Z81506; PIDN:CAB04129.1;

A/Experimental source: clone F16H6

C/Genetics:

A/Gene: CESP:F16H6.3

A/Map position: 5

A/Introns: 39/3; 123/1

C/Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3

Query Match

Best Local Similarity 47.1%; Score 16; DB 2; Length 257;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 195 CTNATTTQASC 206

RESULT 7

T52565 glutathione synthase (BC 6.3.2.3) [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 09-Jul-2004

C/Accession: T52565

R/Skipsey, M.; Andrews, C.J.; Townson, J.K.; Jepson, I.; Edwards, R.

submitted to the EMBL Data Library, July 1999

A/Description: Isolation of cDNA and genomic clones of glutathione synthetase containing

A/Reference number: Z26117

A/Accession: T52565

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-539 <SKI>

A/Cross-references: UNIPROT:P46416; UNIPARC:UPI000016DB47; EMBL:AJ243813; PIDN:CAB51027.

A/Experimental source: cultivar Columbia; seedling; whole plant

C/Genetics:

A/Gene: gsh2

C/Superfamily: glutathione synthase

C/Keywords: ligase

Query Match

Best Local Similarity 47.1%; Score 16; DB 2; Length 539;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 5 CSSLVSSTSC 16

RESULT 8

A48579 trophozoite surface protein TSP11 - Giardia lamblia

C/Species: Giardia lamblia

C/Date: 01-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C/Accession: A48579

R/B, P.L.; Khanna, K.K.; Manning, P.A.; Mayrhofer, G.

Mol. Biochem. Parasitol. 58, 247-257, 1993

A/Title: A gene encoding a 69-kilodalton major surface protein of Giardia intestinalis t:

A/Reference number: A48579; MUID:93241215; PMID:8479449

A/Contents: Ad-1

A/Accession: A48579

A/Status: preliminary

A/Molecule type: nucleic acid

A/Residues: 1-667 <EY1>

A/Cross-references: UNIPROT:Q03185; UNIPARC:UPI0000137726; GB:M95814; NID:9159106; PID:G

A/Note: sequence extracted from NCBI backbone (NCBIN:130056, NCBI:P:130058)

Query Match

Best Local Similarity 47.1%; Score 16; DB 2; Length 667;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

Db 449 CTSTTATVATC 460

RESULT 9

T51024 related to C2H2 zinc finger transcription factor D-Spl [imported] - Neurospora crassa

N/Alternate names: protein B7F21.50

C/Species: Neurospora crassa

C/Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 09-Jul-2004

C/Accession: T51024

R/Schulte, U.; Aign, V.; Hohnsels, J.; Brandt, P.; Fattmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, July 2000

A/Reference number: Z25286

A/Accession: T51024

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-770 <SCH>

A/Cross-references: UNIPROT:Q9P319; UNIPARC:UPI000017B4E0; EMBL:AL389901; GSPDB:GN00116;

A/Experimental source: BAC clone B7F21; strain OR74A

C/Genetics:

A/Gene: NCSP:B7F21.50

A/Map position: 6

A/Introns: 117/1

Query Match

Best Local Similarity 47.1%; Score 16; DB 2; Length 770;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
 Db 92 CSATTAAVGTAC 103

## RESULT 10

T31097  
 chitin synthase (EC 2.4.1.16) Chs1 - fungus (*Filobasidium floriforme*)  
 C/Species: *Filobasidium floriforme*, *Cryptococcus neoformans*  
 C/Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T31097

R/Spectr, C.A.  
 submitted to the EMBL Data Library, August 1997

A/Description: Chs1, a class IV chitin synthase of *Cryptococcus neoformans*.

A/Reference number: Z20980

A/Accession: T31097

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1041 <SPE>

A/Cross-references: UNIPROT:O13356; UNIPARC:UPI00001278CD; EMBL:AF021318; NID:G2444456;

A/Experimental source: strain H99

C/Genetics:

A/Gene: Chs1

A/Introns: 152/1; 556/1; 634/3; 922/2

C/Function:

A/Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucose

A/Pathway: chitin biosynthesis

A/Note: class IV chitin synthase

C/Superfamily: chitin synthase Chs4

C/Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 47.1%; Score 16; DB 2; Length 1041;  
 Best Local Similarity 16.7%; Pred. No. 3.7e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 212 CNSTSSSSGTSC 223

## RESULT 11

AC0416  
 probable prepilin peptidase dependent protein D precursor ppd [imported] - *Yersinia pestis*  
 C/Species: *Yersinia pestis*  
 C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: AC0416

R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.B.

demo-Terraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,

Nature 413, 523-527, 2001

A/Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AC0416

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-154 <KUR>

A/Cross-references: UNIPROT:Q8ZB15; UNIPARC:UPI00000DC98A; GB:AL590842; PIND:CAC92655.1;

C/Genetics:

A/Gene: ppdD

C/Superfamily: gonococcal fimbrin protein

Query Match 44.1%; Score 15; DB 2; Length 154;  
 Best Local Similarity 16.7%; Pred. No. 1.2e-05;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 129 CTATDSAMTSC 140

## RESULT 12

B41378

cytochrome c553i precursor - *Paracoccus denitrificans*

C/Species: *Paracoccus denitrificans*

C/Date: 28-May-1992 #sequence\_revision 28-May-1992 #text\_change 09-Jul-2004

C/Accession: B41378

R/Rae, J.; Reijnders, W.N.M.; Van Spanning, R.J.M.; Harms, N.; Oltmann, L.F.; Stouthamer,

J. Bacteriol. 173, 6971-6979, 1991

A/Title: Isolation, sequencing, and mutagenesis of the gene encoding cytochrome c-553i of

A/Reference number: A41378; MUID:92041583; PMID:1657873

A/Accession: B41378

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-226 <RAS>

A/Cross-references: UNIPROT:P29967; UNIPARC:UPI0000126C5E; GB:M75583; NID:G150576; PIND:J

C/Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-226/Product: cytochrome c553i #status predicted <MAT>

F:115-193/Domain: cytochrome c6 homology <CYC>

F:125,128/Binding site: heme (Cys) (covalent) #status predicted

F:129,173/Binding site: heme iron (His, Met) (axial ligands) #status predicted

Query Match 44.1%; Score 15; DB 2; Length 226;  
 Best Local Similarity 16.7%; Pred. No. 1.3e-05;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 16 CAASAIAGTALC 27

## RESULT 13

T21786

hypothetical protein F35E8.9 - *Caenorhabditis elegans*

C/Species: *Caenorhabditis elegans*

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T21786

R/Gardner, A.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z19472

A/Accession: T21786

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-248 <MTL>

A/Cross-references: UNIPROT:O45450; UNIPARC:UPI0000077E4A; EMBL:Z81529; PIND:CA804292.1;

A/Experimental source: clone F35E8

C/Genetics:

A/Gene: CESP:F35E8.9

A/Map position: 5

A/Introns: 33/3; 121/1

C/Superfamily: *Caenorhabditis elegans* hypothetical protein C49G7.3

Query Match 44.1%; Score 15; DB 2; Length 248;  
 Best Local Similarity 16.7%; Pred. No. 1.3e-05;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15

Db 207 CSSYIADSTAC 218

## RESULT 14

A84746

probable NAM (no apical meristem)-like protein [imported] - *Arabidopsis thaliana*

C/Species: *Arabidopsis thaliana* (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C/Accession: A84746

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.C.

et al. Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: A84746

A/Status: preliminary

A/Molecule type: DNA



A:Residues: 1-268 <STD>  
A:Cross-references: UNIPROT:O22798; UNIPARC:UPI00000A7D54; GB:AE002093; NID:g2459430; PI  
C:Genetics:  
A:Gene: At2g33480  
A:Map position: 2

Query Match 44.1%; Score 15; DB 2; Length 268;  
Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 230 CCSSSTSSSSVC 241

## RESULT 15

S69193  
probable finger protein Pszf1 - garden pea  
C:Species: Pisum sativum (garden pea)  
C>Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S69193  
R:Michael, A.J.; Hofer, J.M.I.; Ellis, T.H.N.  
Plant Mol. Biol. 30, 1051-1058, 1996  
A:Title: Isolation by PCR of a cDNA clone from pea petals with similarity to petunia and  
A:Reference number: S69193; MUID:96270382; PMID:8639742  
A:Accession: S69193  
A:Molecule type: mRNA  
A:Residues: 1-273 <MTC>  
A:Cross-references: UNIPROT:Q41070; UNIPARC:UPI000009E1A7; EMBL:X87374; NID:g854399; PID  
C:Superfamily: Arabidopsis thaliana hypothetical protein F12E4.290  
C:Keywords: DNA binding; zinc finger  
F:115-135/Region: zinc finger CCHH motif  
F:202-222/Region: zinc finger CCHH motif

Query Match 44.1%; Score 15; DB 2; Length 273;  
Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 29 CVANTTVTSSC 40

## RESULT 16

AH2744  
conserved hypothetical protein Atu1368 [imported] - Agrobacterium tumefaciens (strain C5  
C:Species: Agrobacterium tumefaciens  
C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: AH2744  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCell  
; Karp, P.; Kometo, F.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, B.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AH2744  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-283 <KDR>  
A:Cross-references: UNIPROT:Q8UFM9; UNIPARC:UPI00001645DD; GB:AE008688; PIDN:AA142374.1;  
C:Genetics:  
A:Gene: Atu1368  
A:Map position: circular chromosome

Query Match 44.1%; Score 15; DB 2; Length 283;  
Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15

DB 242 CSATSSDMAAC 253

## RESULT 17

S57770  
xyloglucan endo-1,4-beta-D-glucanase (EC 3.2.1.-) precursor (clone tXET-B2) - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C>Date: 06-Dec-1996 #sequence\_revision 07-Feb-1997 #text\_change 09-Jul-2004  
C:Accession: S57770; S49813  
R:Arrowsmith, D.A.; de Silva, J.  
Plant Mol. Biol. 28, 391-403, 1995  
A:Title: Characterisation of two tomato fruit-expressed cDNAs encoding xyloglucan endo-  
A:Reference number: S57769; MUID:95359399; PMID:7632911  
A:Accession: S57770  
A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA  
A:Residues: 1-287 <ARR>  
A:Cross-references: UNIPROT:Q43528; UNIPARC:UPI00000AAE06; EMBL:X82684; NID:g577067; PID  
A>Note: the sequence from Fig. 2 is inconsistent with that from Fig. 5 in having 246-Glu  
C:Superfamily: endoxyloglucan transferase  
C:Keywords: glycosidase; hydrolase  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-287/Product: xyloglucan endo-transglycosylase #status predicted <MAT>

Query Match 44.1%; Score 15; DB 2; Length 287;  
Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 217 CWSAATSTSSC 228

## RESULT 18

S49812  
xyloglucan endo-1,4-beta-D-glucanase (EC 3.2.1.-) precursor (clone tXET-B1) - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C>Date: 05-Mar-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
C:Accession: S57769; S49812  
R:Arrowsmith, D.A.; de Silva, J.  
Plant Mol. Biol. 28, 391-403, 1995  
A:Title: Characterisation of two tomato fruit-expressed cDNAs encoding xyloglucan endo-  
A:Reference number: S57769; MUID:95359399; PMID:7632911  
A:Accession: S57769

A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-289 <ARR>  
A:Cross-references: UNIPROT:Q43527; UNIPARC:UPI00000A4E68; EMBL:X82685; NID:g577065; PID  
C:Genetics:  
A:Gene: tXET-B1  
C:Superfamily: endoxyloglucan transferase  
C:Keywords: glycosidase; hydrolase

Query Match 44.1%; Score 15; DB 2; Length 289;  
Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 219 CWSAASSTSSC 230

## RESULT 19

G97525  
CG10208 protein (AE003744) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)  
C:Species: Agrobacterium tumefaciens  
C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: G97525  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, Y.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: A97359; MUID:21608551; PMID:11743194

A/Accession: G97525  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-318 <KIR>  
A/Cross-references: UNIPROT:Q8UFM9, UNIPARC:UPI00000D1B27, GB:AE007869, PIDN:AAK87160.1,  
C/Genetics:  
A/Gene: AGR\_C\_2528  
A/Map position: circular chromosome

Query Match 44.1%; Score 15; DB 2; Length 318;  
Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 277 CSATHSSDMAAC 288

RESULT 20  
T21152  
hypothetical protein F20E11.7 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T21152  
R/Ainscough, R.  
submitted to the EMBL Data Library, November 1996  
A/Reference number: Z19383  
A/Accession: T21152  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-357 <WIL>  
A/Cross-references: UNIPROT:Q9YV87, UNIPARC:UPI0000060FAD, EMBL:Z81508, PIDN:CAB04138.1,  
A/Experimental source: clone F20E11  
C/Genetics:  
A/Gene: CBSP:F20E11.7  
A/Map position: 5  
A/Intons: 52/3; 88/2; 103/3; 154/3; 190/2; 205/3; 241/2; 270/3; 321/3

Query Match 44.1%; Score 15; DB 2; Length 357;  
Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 7 CTATTGTNSAGC 18

RESULT 21  
S18765  
sd protein - fruit fly (Drosophila melanogaster)  
C/Species: Drosophila melanogaster  
C/Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C/Accession: S18765  
R/Powers, P.A.; Ganetzky, B.  
Genetics 129, 133-144, 1991  
A/Title: On the components of segregation distortion in Drosophila melanogaster. V. Mole  
A/Reference number: S18765; MUID:92038937, PMID:1936954  
A/Accession: S18765  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-363 <PCM>  
A/Cross-references: UNIPROT:P25722, UNIPARC:UPI000016BD5D, EMBL:X60218, NID:g11012; PIDN  
C/Genetics:  
A/Gene: FlyBase:sd  
A/Cross-references: FlyBase:FBgn0024230

Query Match 44.1%; Score 15; DB 2; Length 363;  
Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 331 CSRTSTFRSSC 342

RESULT 22  
T27303  
hypothetical protein Y68A4A.10 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T27303  
R/Steward, C.  
submitted to the EMBL Data Library, January 1998  
A/Reference number: Z20340  
A/Accession: T27303  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-392 <WIL>  
A/Cross-references: UNIPROT:Q9XHP6, UNIPARC:UPI000007BED1, EMBL:AL021503, PIDN:CAA16424.1  
A/Experimental source: clone Y68A4A  
C/Genetics:  
A/Gene: CBSP:Y68A4A.10  
A/Map position: 5  
A/Intons: 45/2; 74/3; 139/3; 175/2; 190/3; 241/3; 305/3; 356/3

Query Match 44.1%; Score 15; DB 2; Length 392;  
Best Local Similarity 16.7%; Pred. No. 1.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 260 CTATTGMTSTGC 271

RESULT 23  
T08928  
sucrose cleavage protein homolog T15N24.70 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C/Accession: T08928  
R/Beyan, M.; Zimmermann, W.; Gruenisen, A.; Wandt, R.; Bancroft, I.; Mewes, H.W.; Maye  
submitted to the Protein Sequence Database, May 1999  
A/Reference number: Z16518  
A/Accession: T08928  
A/Molecule type: DNA  
A/Residues: 1-409 <BEV>  
A/Cross-references: UNIPROT:Q9UUA2, UNIPARC:UPI000009E309, EMBL:AL078465, GSPDB:GN00062;  
A/Experimental source: cultivar Columbia; BMC clone T15N24  
C/Genetics:  
A/Gene: ATSP:T15N24.70  
A/Map position: 4  
A/Intons: 126/3; 155/2; 223/1; 299/3

Query Match 44.1%; Score 15; DB 2; Length 409;  
Best Local Similarity 16.7%; Pred. No. 1.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 394 CAASVAVAYTC 405

RESULT 24  
S74388  
gamma-glutamyl phosphate reductase proA - Synechocystis sp. (strain PCC 6803)  
N/Alternate names: hypothetical protein sl10461  
C/Species: Synechocystis sp.  
A/Variety: PCC 6803  
C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C/Accession: S74388  
R/Kaneke, T.; Sato, S.; Kozani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
o, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,  
DNA Res. 3, 109-136, 1996  
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
s.  
A/Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S74388  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-420 <KAN>  
A:Cross-references: UNIPROT:Q55167; UNIPARC:UPI0000139848; EMBL:664001; GB:AB001339; NID:96718  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: proA  
C:Superfamily: gamma-glutamyl phosphate reductase

Query Match 44.1%; Score 15; DB 2; Length 420;  
Best Local Similarity 16.7%; Pred. No. 1.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15  
Db 153 CGASHSSTAAIC 164

RESULT 25  
Unknown protein, 74043-75895 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: D96552  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizart, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltli, R.; Marzilli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:1130712  
A:Accession: D96552  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-423 <STO>  
A:Cross-references: UNIPROT:Q9C8J4; UNIPARC:UPI00000A997F; GB:AE005173; NID:91092345; F  
C:Genetics:  
A:Gene: FSD21.9  
A:Map position: 1  
C:Superfamily: Arabidopsis thaliana hypothetical protein F24B22.230

Query Match 44.1%; Score 15; DB 2; Length 423;  
Best Local Similarity 16.7%; Pred. No. 1.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15  
Db 306 CTTYTSILKASC 317

RESULT 26  
S31126  
hypothetical protein F59B2.6 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S31126  
R:Stultson, J.; Du, Z.; Thomas, K.; Wilson, R.; Hillier, L.; Staden, R.; Halloran, N.; Gr awkins, T.; Alnough, R.; Waterston, R.  
submitted to the EMBL Data Library, November 1991  
A:Description: The C. elegans sequencing project: A beginning.  
A:Reference number: S31122  
A:Accession: S31126  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-503 <STU>  
A:Cross-references: UNIPROT:P34482; UNIPARC:UPI00001399DD; EMBL:Z11505; NID:96718; PID:9  
C:Genetics:  
A:Introns: 26/3; 42/2; 121/3; 367/3; 443/1

C:Superfamily: Caenorhabditis elegans hypothetical protein F59B2.6

Query Match 44.1%; Score 15; DB 2; Length 503;  
Best Local Similarity 16.7%; Pred. No. 1.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15  
Db 4 CSASTSQLSTFC 15

RESULT 27  
B87180  
pyruvate (or indolepyruvate) decarboxylase [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 05-Oct-2004  
C:Accession: B87180  
R:Coile, S.T.; Bigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho ram, M.A.; Rutherford, K.M.  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
A:Title: Massive gene decay in the leprosy bacillus.  
A:Reference number: A86909; MUID:21128732; PMID:11234002  
A:Accession: B87180  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-569 <STO>  
A:Cross-references: UNIPROT:Q9C8D6; UNIPARC:UPI00000C6E3D; GB:AL450380; NID:913093715; P  
C:Genetics:  
A:Gene: pdc  
C:Superfamily: pyruvate decarboxylase/indolepyruvate decarboxylase; thiamin pyrophosphat

Query Match 44.1%; Score 15; DB 2; Length 569;  
Best Local Similarity 16.7%; Pred. No. 1.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15  
Db 140 CSQANLTTATAC 151

RESULT 28  
JC7818  
metalloprotease I precursor - Alteromonas sp. (Strain O-7)  
C:Species: Alteromonas sp. (Strain O-7)  
C:Date: 03-Jun-2002 #sequence\_revision 03-Jun-2002 #text\_change 31-Dec-2004  
C:Accession: JC7818  
R:Myamoto, K.; Tsujibo, H.; Nukui, E.; Itoh, H.; Kaidzu, Y.; Inamori, Y.  
Biosci. Biotechnol. Biochem. 66, 416-421, 2002  
A:Title: Isolation and characterization of the genes encoding two metalloproteases (Mpri  
A:Reference number: JC7818; PMID:1199419; MUID:21994062  
A:Accession: JC7818  
A:Molecule type: DNA  
A:Residues: 1-727 <MTY>  
A:Cross-references: UNIPROT:Q8VLO5; UNIPARC:UPI000017CD49; DDBJ:AB063611  
A:Experimental source: strain O-7  
C:Comment: This enzyme, an extracellular alkaline metalloprotease, showing high similar  
C:Genetics:  
A:Gene: mpri  
A:Start codon: GTG  
C:Superfamily: zinc metalloendopeptidase, neutral protease type (elastase)

Query Match 44.1%; Score 15; DB 2; Length 727;  
Best Local Similarity 16.7%; Pred. No. 1.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15  
Db 581 CNSTTSNSC 592

RESULT 29



A:Gene: ATSP:T11111.140  
A:Map position: 4  
A:Insertions: 35/1; 66/2; 115/3; 613/3; 694/3; 799/3; 879/3; 974/3; 1048/3; 1102/3; 1147/3;  
C:Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology  
C:Keywords: 2Fe-2S; metalloprotein  
F:43,48,51,86/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 44.1%; Score 15; DB 2; Length 1364;  
Best Local Similarity 16.7%; Pred. No. 1.9e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
Db 209 CGSKTSEATRC 220

RESULT 34  
S65571  
Pattern formation protein GNOM - Arabidopsis thaliana

N:Alternate names: EMB30 protein  
C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 09-Jul-2004  
C:Accession: S65571; S65572

R:Busch, M.; Mayer, U.; Juegens, G.  
Mol. Gen. Genet. 250, 681-691, 1996

A:Title: Molecular analysis of the Arabidopsis pattern formation gene GNOM: gene structure  
A:Reference number: S65571; MUID:96204508; PMID:8628228

A:Accession: S65571  
A:Molecule type: DNA

A:Residues: 1-1451 <BUS>  
A:Cross-references: UNIPROT:Q42510; UNIPARC:UPI0000129B82; EMBL:U36433; NID:g1209632; PI

A:Accession: S65572  
A:Molecule type: mRNA

A:Status: nucleic acid sequence not shown; translation not shown  
A:Residues: 1-110; 'T', 112-866, 'G', 868-1451 <BUS>

A:Cross-references: UNIPARC:UPI000016DA45; EMBL:U36432; NID:g1209630; PID:g1209631  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995

C:Genetics:  
A:Gene: GNOM; EMB30

A:Introns: 246/3

Query Match 44.1%; Score 15; DB 2; Length 1451;  
Best Local Similarity 16.7%; Pred. No. 1.9e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
Db 22 CTSSNTTTLAC 33

RESULT 35  
T25888  
hypothetical protein T10H10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T25888

R:Nelson, J.; Langston, Y.  
submitted to the EMBL Data Library, December 1996

A:Description: The sequence of C. elegans cosmid T10H10.  
A:Reference number: Z20105

A:Accession: T25888  
A:Molecule type: DNA

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Residues: 1-2098 <NEL>

A:Cross-references: UNIPROT:P91443; UNIPARC:UPI00000756CE; EMBL:U80848; PIDN:AAB37988.1;  
A:Experimental source: strain Bristol N2; clone T10H10

C:Genetics:  
A:Gene: CESP:T10H10.1

A:Map position: X  
A:Insertions: 6/3; 43/3; 92/3; 154/2; 226/1; 280/3; 417/2; 472/3; 812/3; 853/3; 887/3; 978/  
/3

F:65-720/Domain: myosin motor domain homology <MMO>

Query Match 44.1%; Score 15; DB 2; Length 2098;  
Best Local Similarity 16.7%; Pred. No. 2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
Db 1188 CADSATTAELC 1199

RESULT 36

T13806  
toucan gene protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster  
C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004

C:Accession: T13806  
R:Grammont, M.; Dastugue, B.; Couderc, J.L.

Development 124, 4917-4926, 1997  
A:Title: The Drosophila toucan (tcc) gene is required in germline cells for somatic cell

A:Reference number: Z17769; MUID:98090047; PMID:9362455  
A:Accession: T13806

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA

A:Residues: 1-2176 <GRA>  
A:Cross-references: UNIPROT:O46112; UNIPARC:UPI0000083F2C; EMBL:Y14157; NID:g2760521; PI

C:Genetics:  
A:Gene: tcc  
A:Cross-references: FlyBase:FBgn0015600

Query Match 44.1%; Score 15; DB 2; Length 2176;  
Best Local Similarity 16.7%; Pred. No. 2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
Db 2031 CIRASATSAVC 2042

RESULT 37  
S60224

polyketide synthase 1 - Colletotrichum lagenarium

C:Species: Colletotrichum lagenarium  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004

C:Accession: S60224  
R:Takano, Y.; Kubo, Y.; Shimizu, K.; Mise, K.; Okuno, T.; Furusawa, I.

Mol. Gen. Genet. 249, 162-167, 1995  
A:Title: Structural analysis of PKS1, a polyketide synthase gene involved in melanin bios

A:Reference number: S60224; MUID:96086932; PMID:7500937  
A:Accession: S60224

A:Status: not compared with conceptual translation  
A:Molecule type: DNA

A:Residues: 1-2187 <TAK>  
A:Cross-references: UNIPROT:P79068; UNIPARC:UPI000006C308; GB:D83643; NID:g1208941; PIDN:

C:Genetics:  
A:Gene: pks1  
A:Function:

A:Pathway: melanin biosynthesis  
C:Keywords: carrier protein

F:402-811/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>  
F:910-1198/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>

F:166-1736/Domain: acyl carrier protein homology <ACP>  
F:1788-1862/Domain: acyl carrier protein homology <ACP1>

Query Match 44.1%; Score 15; DB 2; Length 2187;  
Best Local Similarity 16.7%; Pred. No. 2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
Db 553 CSSSAAALNVAC 564

RESULT 38  
T28675

alpha-51D immobilization antigen - Paramedium tetraurelia  
C/Species: Paramedium tetraurelia  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T28675  
R/Schwegmann, K.J.  
Submitted to the EMBL Data Library, March 1996  
A/Reference number: Z20506  
A/Accession: T28675  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-2533 <SCH>  
A/Cross-references: UNIPROT:P90589, UNIPARC:UPI0000078A73, EMBL:X96400, PIDN:CA65264.1  
C/Genetics:  
A/Genetic code: SGC5  
A/Intons: 280/3; 538/2; 1248/2  
C/Superfamily: G surface protein

Query Match 44.1%; Score 15; DB 2; Length 2533;  
Best Local Similarity 16.7%; Pred. No. 2.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
Db 360 CKGTSTTEAAC 371

RESULT 39  
T28674  
alpha-51D-immobilization antigen - Paramedium tetraurelia  
C/Species: Paramedium tetraurelia  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T28674  
R/Schmidt, H.J.  
submitted to the EMBL Data Library, March 1995  
A/Reference number: Z20505  
A/Accession: T28674  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-2533 <SCH>  
A/Cross-references: UNIPROT:Q27183, UNIPARC:UPI000007E3B2, EMBL:X85135, NID:g128634, PID  
C/Genetics:  
A/Genetic code: SGC5  
A/Note: alpha-51D  
C/Superfamily: G surface protein

Query Match 44.1%; Score 15; DB 2; Length 2533;  
Best Local Similarity 16.7%; Pred. No. 2.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
Db 360 CKGTSTTEAAC 371

RESULT 40  
T31687  
surface antigen - Paramedium primaurelia  
C/Species: Paramedium primaurelia  
C/Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C/Accession: T31687  
R/Bourgain-Guglielmetti, F.; Caron, F.  
Journal of Eukaryot. Microbiol. 43, 303-314, 1996  
A/Title: Molecular characterization of the D surface protein gene subfamily in Paramedium  
A/Reference number: Z21061, MUID:56313351, PMID:8768434  
A/Accession: T31687  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-2543 <BOU>  
A/Cross-references: UNIPROT:P90649, UNIPARC:UPI000007A903, EMBL:X96616, NID:g1235576, PID  
C/Genetic code: SGC5  
A/Genetic family: G surface protein

Query Match 44.1%; Score 15; DB 2; Length 2543;  
Best Local Similarity 16.7%; Pred. No. 2.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
Db 360 CKGTSTTEASC 371

RESULT 41  
A24420  
notch protein - fruit fly (Drosophila melanogaster)  
N/Alternate names: neurogenic repetitive locus protein  
C/Species: Drosophila melanogaster  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 05-Oct-2004  
C/Accession: A24420, A24768, S09358, A05267  
R/Kidd, S.; Kelley, M.R.; Young, M.W.  
Mol. Cell. Biol. 6, 3094-3108, 1986  
A/Reference number: A24420, MUID:87064624, PMID:3097517  
A/Accession: A24420  
A/Molecule type: DNA  
A/Residues: 1-2703 <KID>  
A/Cross-references: UNIPROT:P07207, UNIPARC:UPI000016SC66, GB:X03508, NID:g157991, PIDN:  
R/Marton, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S.  
Cell 43, 567-581, 1985  
A/Reference number: A24768, MUID:86079539, PMID:3935325  
A/Accession: A24768  
A/Molecule type: mRNA  
A/Residues: 1-48, 'T', 50-118, 'R', 120-230, 'T', 232-256, 'N', 258-266, 'A', 268-872, 'R', 874-958, '  
A/Cross-references: UNIPARC:UPI0000173D1F  
A/Note: the authors translated the codon ATC for residue 49 as Thr, ATT for residue 2044  
R/Tautz, D.  
Nucleic Acids Res. 17, 6463-6471, 1989  
A/Title: Hypervariability of simple sequences as a general source for polymorphic DNA mar  
A/Reference number: S09358, MUID:89385974, PMID:2780284  
A/Accession: S09358  
A/Molecule type: DNA  
A/Residues: 2505-2551, 'OQOO', 2552-2576, 'B', 2578-2604 <TAU>  
A/Cross-references: UNIPARC:UPI0000173D20  
R/Marton, K.A.; Yedobnick, B.; Flimerty, V.G.; Artavanis-Tsakonas, S.  
Cell 40, 55-62, 1985  
A/Title: opa: a novel family of transcribed repeats shared by the Notch locus and other c  
A/Reference number: A05267, MUID:85099329, PMID:2981631  
A/Accession: A05267  
A/Molecule type: DNA  
A/Residues: 2504-2576, 'E', 2578-2611 <WHA2>  
A/Cross-references: UNIPARC:UPI0000173D21  
C/Genetics:  
A/Genetic code: SGC5  
A/Genetic family: G surface protein  
A/Map position: 8.96-9.36  
A/Intons: 53/3; 84/3; 171/3; 240/3; 283/3; 2333/3; 2436/3; 2588/3  
C/Superfamily: notch protein; ankyrin repeat homology; EGF homology  
C/Keywords: differentiation; tandem repeat; transmembrane protein  
F/27-43/Domain: transmembrane #status predicted <TMW1>  
F/297-328/Domain: EGF homology <EGX1>  
F/530-561/Domain: EGF homology <EGF1>  
F/568-599/Domain: EGF homology <EGF>  
F/968-1019/Domain: EGF homology <EGX2>  
F/1064-1095/Domain: EGF homology <EGX3>  
F/1187-1218/Domain: EGF homology <EGX3>  
F/1746-1765/Domain: transmembrane #status predicted <TMW2>  
F/1950-1989/Domain: ankyrin repeat homology <AN1>  
F/1983-2015/Domain: ankyrin repeat homology <AN2>  
F/1988-2004/Domain: transmembrane #status predicted <TMW3>  
F/2017-2049/Domain: ankyrin repeat homology <AN3>  
F/2050-2082/Domain: ankyrin repeat homology <AN4>  
F/2083-2115/Domain: ankyrin repeat homology <AN5>  
F/2538-2568/Region: glutamine-rich  
F/2538-2568/Domain: neurogenic repetitive element #status predicted <OPA>

Query Match 44.1%; Score 15; DB 1; Length 2703;

Best Local Similarity 16.7%; Pred. No. 2.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 192 CTALAGSSSFTC 203

## RESULT 42

hypothetical protein G01D9.5 - Caenorhabditis briggsae

C:Species: Caenorhabditis briggsae

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T11345

R:Waterston, R.

submitted to the EMBL Data Library, April 1996

A:Description: The C. briggsae genome sequencing project.

A:Reference number: 221010

A:Accession: T11345

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4767 <MAT>

A:Cross-references: UNIPROT:Q17301, UNIPARC:UPI000011013B, EMBL:U56248, NID:g1293789, PI

C:Genetics:

A:Insertions: 269/3; 341/3; 853/1; 920/2; 4452/3; 4534/3; 4592/3; 4654/3; 4670/3; 4707/2; 4

A:Note: G01D9.5

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein

F:3472-3541/Domain: acyl carrier protein homology <ACPL>

F:4039-4427/Domain: acetate-CoA ligase homology <ACLI>

F:4447-4514/Domain: acyl carrier protein homology <ACPL>

F:2210.3505/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match

Best Local Similarity 44.1%; Score 15; DB 2; Length 4767;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 1423 CSSSFHALSSAC 1434

## RESULT 43

hypothetical protein C41A3.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C:Accession: T15789

R:Bentley, D.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C41A3.

A:Reference number: 218404

A:Accession: T15789

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-7829 <BEN>

A:Cross-references: UNIPROT:Q18559, UNIPARC:UPI000011014C, EMBL:U41541, NID:g1109867, PI

C:Genetics:

A:Gene: C41A3.1

A:Insertions: 29/2; 69/3; 141/2; 192/2; 271/3; 321/2; 1000/1; 1098/1; 1127/3; 1168/2; 1282/

2; 7514/3; 7556/3; 7654/3; 7716/3; 7732/3; 7769/2; 7797/3

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein

F:6547-6616/Domain: acyl carrier protein homology <ACP>

F:2832.5271.6580/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match

Best Local Similarity 44.1%; Score 15; DB 2; Length 7829;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 4475 CSSSFYALSSAC 4486

RESULT 44  
118346  
elastic titin - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004

C:Accession: I18346

R:Labell, S.; Kolmerer, B.

A:Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.

A:Reference number: A57430; MUID:96026330; PMID:7569978

A:Accession: I18346

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-7962 <RES>

A:Cross-references: UNIPROT:Q10465, UNIPARC:UPI000011010C, EMBL:X90569, NID:g1017426, PI

C:Genetics:

A:Gene: GDB:TTN

A:Cross-references: GDB:127867; OMIM:188840

A:Map position: 2q31-2q31

Query Match

Best Local Similarity 44.1%; Score 15; DB 2; Length 7962;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 381 CTASNEAGSSSC 392

## RESULT 45

R-phycocerythrin beta-2 chain - red alga (Gastrocoulonium coulteri) (fragment)

C:Species: Gastrocoulonium coulteri

C:Date: 07-Mar-1998 #sequence\_revision 07-Mar-1998 #text\_change 09-Jul-2004

C:Accession: D22565

R:Klotz, A.V.; Glazer, A.N.

J. Biol. Chem. 260, 4856-4863, 1985

A:Title: Characterization of the bilin attachment sites in R-phycocerythrin.

A:Reference number: A22565; MUID:95182601; PMID:3886644

A:Accession: D22565

A:Molecule type: protein

A:Residues: 1-21 <KLO>

A:Cross-references: UNIPROT:Q7M276; UNIPARC:UPI0000174E0D

C:Superfamily: phycocyanin

Query Match

Best Local Similarity 41.2%; Score 14; DB 2; Length 21;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 8 CSMSSSEVASYC 19

## RESULT 46

hypothetical protein XF0861 [imported] - Xylella fastidiosa (strain 9asc)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C:Accession: G82754

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10510347

A:Note: For a complete list of authors see reference number A59328 below

A:Accession: G82754

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-88 <SIM>

A:Cross-references: UNIPROT:Q9PF17, UNIPARC:UPI00000C254B; GB:AE003925; GB:AE003849; NID:

A:Experimental source: strain 9asc

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carre, H

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincant, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; From  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitejima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
Rodrigues, V.; Rosa, A.J. de M.; de M.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, W.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
A:Genetics:  
A:Gene: XF0861

Query Match 41.2%; Score 14; DB 2; Length 88;  
Best Local Similarity 16.7%; Pred. No. 0.00052;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 23 CXXSVTLKSSRC 34

#### RESULT 47

154781 fibroblast growth factor receptor FGFR-1, secreted splice form - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 21-Jul-2000

C:Accession: 154781

R:Leung, H.Y.; Gullick, W.J.; Lemoine, N.R.

Int. J. Cancer 59, 667-675, 1994

A:Title: Expression and functional activity of fibroblast growth factors and their recep

A:Accession: 154781

A:Reference number: 154781; MUID:95048906; PMID:7960240

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-92 <RES>

A:Cross-references: UNIPARC:UPI0000116DB9; GB:S76658; NID:913557; PIDN:AMD14230.1; PID:

A:Experimental source: pancreatic cancer cell line PT45

C:Genetics:

A:Gene: GDB:FGFR1, FLT2

A:Cross-references: GDB:119913; OMIM:136350

A:Map position: Bp11.2-Bp11.1

A:Insertions: 58/3

C:Keywords: growth factor receptor

Query Match 41.2%; Score 14; DB 2; Length 92;  
Best Local Similarity 16.7%; Pred. No. 0.00053;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 30 CGAGSASSGRC 41

#### RESULT 48

T28947 hypothetical protein F07C4.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T28947

R:Miller, N.; Stellyes, L.

submitted to the EMBL Data Library, January 1997

A:Description: The sequence of C. elegans cosmid F07C4.

A:Reference number: Z20546

A:Accession: T28947

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-97 <MTL>

A:Cross-references: UNIPROT:P91216; UNIPARC:UPI00000737D; EMBL:U80023; PIDN:AAC48019.1;

A:Experimental source: strain Bristol N2; clone F07C4

C:Genetics:

A:Gene: CESP:F07C4.11

A:Map position: 5

Query Match 41.2%; Score 14; DB 2; Length 97;  
Best Local Similarity 16.7%; Pred. No. 0.00053;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 56 CTTTAAADSPSC 67

#### RESULT 49

C72775 hypothetical protein APE0189 - Aeropyrum pernix (strain KI)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C:Accession: C72775

R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: C72775

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-100 <KAM>

A:Cross-references: UNIPROT:Q9YFRO; UNIPARC:UPI000005DA3D; DDBJ:AP000058; NID:95103388; I

A:Experimental source: strain KI

C:Genetics:

A:Gene: APE0189

C:Superfamily: Aeropyrum pernix hypothetical protein APE0189

Query Match 41.2%; Score 14; DB 2; Length 100;  
Best Local Similarity 16.7%; Pred. No. 0.00054;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 65 CASVSRMARASC 76

#### RESULT 50

T14780 hypothetical protein DKFZp566B1346.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T14780

R:Ostenmaier, B.; Obermaier, B.; Mewes, H.W.; Gaassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, August 1999

A:Reference number: Z18184

A:Accession: T14780

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-101 <OTT>

A:Cross-references: UNIPARC:UPI000016AC6B; EMBL:AL110253

A:Experimental source: fetal kidney; clone DKFZp566B1346

C:Genetics:

A>Note: DKFZp566B1346.1

Query Match 41.2%; Score 14; DB 2; Length 101;  
Best Local Similarity 16.7%; Pred. No. 0.00054;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
DB 38 CTSISVVTQTASC 49

Search completed: January 4, 2006, 16:10:25  
Job time : 26.6261 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 : Search time 117.548 Seconds  
(without alignments)  
108.037 Million cell updates/sec

Title: US-09-932-322-5  
Perfect score: 34  
Sequence: 1 XXXXXXXXXXXXXXXX 18

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : Uniprot 05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	52.9	307	1	SGS3_DROME
2	18	52.9	347	2	OSK340_MOUSE
3	18	52.9	362	2	OSK303_MOUSE
4	18	52.9	367	2	OSK308_MOUSE
5	18	52.9	384	2	OSK308_MOUSE
6	18	52.9	730	2	OSK308_MOUSE
7	17	50.0	35	2	OSK308_MOUSE
8	17	50.0	137	2	OSK308_MOUSE
9	17	50.0	137	2	OSK308_MOUSE
10	17	50.0	210	2	OSK308_MOUSE
11	17	50.0	234	2	OSK308_MOUSE
12	17	50.0	245	2	OSK308_MOUSE
13	17	50.0	306	2	OSK308_MOUSE
14	17	50.0	378	2	OSK308_MOUSE
15	17	50.0	396	2	OSK308_MOUSE
16	17	50.0	429	2	OSK308_MOUSE
17	17	50.0	540	2	OSK308_MOUSE
18	17	50.0	667	2	OSK308_MOUSE
19	17	50.0	1031	2	OSK308_MOUSE
20	17	50.0	5703	1	MUC5B_HUMAN
21	16	47.1	93	2	OSK308_MOUSE
22	16	47.1	100	2	OSK308_MOUSE
23	16	47.1	100	2	OSK308_MOUSE
24	16	47.1	110	1	CBP1_ADSCA
25	16	47.1	144	2	OSK308_MOUSE
26	16	47.1	152	2	OSK308_MOUSE
27	16	47.1	161	2	OSK308_MOUSE
28	16	47.1	172	2	OSK308_MOUSE
29	16	47.1	176	2	OSK308_MOUSE
30	16	47.1	187	2	OSK308_MOUSE
31	16	47.1	208	2	OSK308_MOUSE

32	16	47.1	210	2	OSK308_MOUSE
33	16	47.1	214	2	OSK308_MOUSE
34	16	47.1	215	2	OSK308_MOUSE
35	16	47.1	215	2	OSK308_MOUSE
36	16	47.1	227	2	OSK308_MOUSE
37	16	47.1	230	2	OSK308_MOUSE
38	16	47.1	234	2	OSK308_MOUSE
39	16	47.1	234	2	OSK308_MOUSE
40	16	47.1	244	2	OSK308_MOUSE
41	16	47.1	247	2	OSK308_MOUSE
42	16	47.1	248	2	OSK308_MOUSE
43	16	47.1	257	2	OSK308_MOUSE
44	16	47.1	262	2	OSK308_MOUSE
45	16	47.1	269	2	OSK308_MOUSE
46	16	47.1	278	2	OSK308_MOUSE
47	16	47.1	285	2	OSK308_MOUSE
48	16	47.1	286	2	OSK308_MOUSE
49	16	47.1	287	2	OSK308_MOUSE
50	16	47.1	297	2	OSK308_MOUSE
51	16	47.1	305	2	OSK308_MOUSE
52	16	47.1	309	2	OSK308_MOUSE
53	16	47.1	320	2	OSK308_MOUSE
54	16	47.1	325	2	OSK308_MOUSE
55	16	47.1	329	2	OSK308_MOUSE
56	16	47.1	332	2	OSK308_MOUSE
57	16	47.1	353	2	OSK308_MOUSE
58	16	47.1	365	2	OSK308_MOUSE
59	16	47.1	381	2	OSK308_MOUSE
60	16	47.1	381	2	OSK308_MOUSE
61	16	47.1	426	2	OSK308_MOUSE
62	16	47.1	437	2	OSK308_MOUSE
63	16	47.1	448	2	OSK308_MOUSE
64	16	47.1	465	2	OSK308_MOUSE
65	16	47.1	471	2	OSK308_MOUSE
66	16	47.1	499	2	OSK308_MOUSE
67	16	47.1	504	2	OSK308_MOUSE
68	16	47.1	534	2	OSK308_MOUSE
69	16	47.1	538	2	OSK308_MOUSE
70	16	47.1	539	1	OSK308_MOUSE
71	16	47.1	543	2	OSK308_MOUSE
72	16	47.1	544	2	OSK308_MOUSE
73	16	47.1	556	2	OSK308_MOUSE
74	16	47.1	569	2	OSK308_MOUSE
75	16	47.1	622	2	OSK308_MOUSE
76	16	47.1	667	1	OSK308_MOUSE
77	16	47.1	709	2	OSK308_MOUSE
78	16	47.1	717	2	OSK308_MOUSE
79	16	47.1	719	2	OSK308_MOUSE
80	16	47.1	719	2	OSK308_MOUSE
81	16	47.1	729	2	OSK308_MOUSE
82	16	47.1	760	2	OSK308_MOUSE
83	16	47.1	772	2	OSK308_MOUSE
84	16	47.1	776	1	OSK308_MOUSE
85	16	47.1	776	1	OSK308_MOUSE
86	16	47.1	776	1	OSK308_MOUSE
87	16	47.1	799	2	OSK308_MOUSE
88	16	47.1	868	2	OSK308_MOUSE
89	16	47.1	1041	1	OSK308_MOUSE
90	16	47.1	1046	2	OSK308_MOUSE
91	16	47.1	1089	2	OSK308_MOUSE
92	16	47.1	1107	2	OSK308_MOUSE
93	16	47.1	1136	2	OSK308_MOUSE
94	16	47.1	1167	2	OSK308_MOUSE
95	16	47.1	1178	2	OSK308_MOUSE
96	16	47.1	1236	2	OSK308_MOUSE
97	16	47.1	1341	2	OSK308_MOUSE
98	16	47.1	1494	2	OSK308_MOUSE
99	16	47.1	1519	2	OSK308_MOUSE
100	16	47.1	1592	2	OSK308_MOUSE
101	16	47.1	5079	2	OSK308_MOUSE
102	16	47.1	35	2	OSK308_MOUSE
103	16	47.1	36	2	OSK308_MOUSE
104	16	47.1	43	2	OSK308_MOUSE

105	15	44.1	57	2	094MN5_9CAUD	094mns bacterioph	178	15	44.1	227	2	09VA83_DROME	09va83 dirosophila
106	15	44.1	57	2	04MOU9_BACCE	04mu9 bacillus ce	179	15	44.1	229	2	07Z969_9FUNG	07z969 fungal endo
107	15	44.1	57	2	04MOV1_BACCE	04mov1 bacillus ce	180	15	44.1	229	2	07Z973_9FUNG	07z973 fungal endo
108	15	44.1	72	2	05CAR3_SCHUA	05car3 schistosoma	181	15	44.1	229	2	07Z9K3_9FUNG	07z9K3 fungal endo
109	15	44.1	79	2	04MOV6_BACCE	04mov6 bacillus ce	182	15	44.1	229	2	07Z9K4_9FUNG	07z9K4 fungal endo
110	15	44.1	86	2	09Y33_9HIV1	09y33 human immun	183	15	44.1	230	2	056P86_9MERT	056p86 fungal endo
111	15	44.1	87	2	07STM4_NEUCR	07stm4 neurospora	184	15	44.1	234	2	056P87_9FUNG	056p87 fungal sp.
112	15	44.1	87	2	04ROQ2_TETNG	04roq2 tetradodon n	185	15	44.1	236	2	056P94_9MERT	056p94 uncultured
113	15	44.1	94	2	06IG47_DROME	06ig47 dirosophila	186	15	44.1	232	2	096UV6_OPHTL	096uv6 ophistoma
114	15	44.1	94	2	09DKL7_VIVIRU	09dkl7 spodoptera	187	15	44.1	242	2	051T39_MAGGR	051t39 magnaportha
115	15	44.1	98	2	07OPR3_GIALA	07opr3 giardia lam	188	15	44.1	243	2	065XA5_ORYSA	065xa5 oryza sativ
116	15	44.1	99	2	06UTV5_CRIPE	06utv5 crinipellis	189	15	44.1	245	2	08FBS9_ECOL6	08fbs9 escherichia
117	15	44.1	99	2	04SDQ4_TETNG	04sdq4 tetradodon n	190	15	44.1	248	2	061HW9_CAEBR	061hw9 caenorhabdi
118	15	44.1	106	2	06TSJ1_PARBR	06tsj1 paracoccidi	191	15	44.1	248	2	045450_CAEBL	045450 caenorhabdi
119	15	44.1	107	2	06H6N9_ORYSA	06h6n9 oryza sativ	192	15	44.1	249	2	0624V6_CAEBR	0624v6 caenorhabdi
120	15	44.1	117	2	06H5N3_ORYSA	06h5n3 oryza sativ	193	15	44.1	250	2	0624V4_CAEBR	0624v4 caenorhabdi
121	15	44.1	125	2	084NA4_9PABA	084na4 dialium gui	194	15	44.1	252	2	061T35_CAEBR	061t35 caenorhabdi
122	15	44.1	128	2	04LIT8_9BURK	04lit8 burkholderi	195	15	44.1	252	2	0624V5_CAEBR	0624v5 caenorhabdi
123	15	44.1	132	2	08R5F9_MOUSE	08r5f9 mus musculu	196	15	44.1	257	2	07PZD8_ANOCA	07pzd8 anopheles g
124	15	44.1	139	2	072A85_DESVH	072a85 desulfovibr	197	15	44.1	257	2	0676U4_9SOLA	0676u4 nictoliana a
125	15	44.1	141	2	0825Q9_STRAW	0825q9 streptomyc	198	15	44.1	261	2	06FPE4_CANGA	06fpe4 candida gla
126	15	44.1	150	2	05Z6U8_ORYSA	05z6u8 oryza sativ	199	15	44.1	264	2	040805_LEIMA	040805 leishmania
127	15	44.1	152	2	06RZ17_9PEZI	06rz17 arthrinium	200	15	44.1	264	2	05D089_9CORO	05d089 canine coro
128	15	44.1	153	2	063T76_BURPS	063t76 burkholderi	201	15	44.1	268	2	022798_ARATH	022798 arabidopsis
129	15	44.1	154	2	066E16_YERPS	066e16 yerania ps	202	15	44.1	269	2	06ZPE5_ORYSA	06zpe5 oryza sativ
130	15	44.1	154	2	08ZB15_YERPE	08zb15 yerania pe	203	15	44.1	273	2	041070_PEA	041070 pisum sativ
131	15	44.1	155	2	07YSR5_DROME	07ysr5 dirosophila	204	15	44.1	281	2	05TX10_ANOCA	05tx10 anocsa
132	15	44.1	155	2	06YIL9_DROYA	06yil9 dirosophila	205	15	44.1	282	2	07RKP8_PLAYO	07rkp8 plasmodium
133	15	44.1	155	2	07ZIU0_DROSI	07ziu0 dirosophila	206	15	44.1	282	2	04LH85_9BURK	04lh85 burkholderi
134	15	44.1	155	2	07ZIU1_DROME	07ziu1 dirosophila	207	15	44.1	283	2	08UTM9_AGRTS	08utm9 agrobacteri
135	15	44.1	155	2	07XEO0_ORYSA	07xEO0 oryza sativ	208	15	44.1	285	2	05LX63_SILPO	05lx63 silpicebaco
136	15	44.1	155	2	093TQ7_9DELT	093tq7 uncultured	209	15	44.1	287	2	043528_LYCES	043528 lycopersico
137	15	44.1	163	2	059LP3_CANAL	059lp3 candida alb	210	15	44.1	289	2	043527_LYCES	043527 lycopersico
138	15	44.1	163	2	09BIM0_GIALA	09bim0 giardia lam	211	15	44.1	293	2	061YT3_CAEBR	061yt3 caenorhabdi
139	15	44.1	163	2	07MOD6_VIBVY	07mqd6 vibrio vuln	212	15	44.1	294	2	09V7R3_DROME	09v7r3 dirosophila
140	15	44.1	165	2	06RZ16_9PEZI	06rz16 apiospora m	213	15	44.1	306	2	07F8S0_ORYSA	07f8s0 oryza sativ
141	15	44.1	165	2	08DD94_VIBVU	08dd94 vibrio vuln	214	15	44.1	318	2	07CZES_AGRTS	07czes agrobacteri
142	15	44.1	169	2	04MHT3_BACCE	04mht3 bacillus ce	215	15	44.1	332	2	062009_CERCA	062009 cercarialia c
143	15	44.1	172	2	04NFB6_9MICC	04nfb6 arthrobacte	216	15	44.1	334	2	08S202_DROME	08s202 dirosophila
144	15	44.1	173	2	07UGY7_RHOBA	07ugy7 rhodospirill	217	15	44.1	334	2	052V10_CIOIN	052v10 ciona intes
145	15	44.1	177	2	05ZIL9_MYCHY	05zil9 mycoplasma	218	15	44.1	334	2	09VW95_DROME	09vw95 dirosophila
146	15	44.1	179	2	04NVY6_9DELT	04nvY6 anaeromyxob	219	15	44.1	335	2	052VH9_CIOIN	052vH9 ciona intes
147	15	44.1	181	2	06IUD0_DROME	06iud0 dirosophila	220	15	44.1	341	2	052V12_CIOIN	052v12 ciona intes
148	15	44.1	183	2	08S3W8_HORVD	08s3w8 hordeum vul	221	15	44.1	341	2	056N31_CIOIN	056n31 ciona intes
149	15	44.1	189	2	082PE7_STRAW	082pe7 streptomyc	222	15	44.1	342	2	053PA6_ORYSA	053pa6 oryza sativ
150	15	44.1	192	2	05SRJ2_CRYNE	05srj2 cryptococcu	223	15	44.1	343	2	09IFH6_9INFA	09ifh6 influenza a
151	15	44.1	192	2	05KG00_CRYNE	05kg00 cryptococcu	224	15	44.1	343	2	09IFH7_9INFA	09ifh7 influenza a
152	15	44.1	193	2	07XIH6_ORYSA	07xih6 oryza sativ	225	15	44.1	343	2	04TDD4_TETNG	04tdD4 tetradodon n
153	15	44.1	195	2	07QOG1_GIALA	07qog1 giardia lam	226	15	44.1	347	2	08STW5_DROME	08stw5 dirosophila
154	15	44.1	196	2	053M55_ORYSA	053m55 oryza sativ	227	15	44.1	347	2	08G4A1_BIFLO	08g4a1 bifidobacte
155	15	44.1	199	2	06IG98_DROME	06ig98 dirosophila	228	15	44.1	350	2	06Z8S2_ORYSA	06z8s2 oryza sativ
156	15	44.1	209	2	04QSG6_9LECA	04qsg6 pertusaria	229	15	44.1	352	2	04S7B8_TETNG	04s7b8 tetradodon n
157	15	44.1	213	2	06V282_9BACT	06v282 symbiont ba	230	15	44.1	354	2	06P898_XENTR	06p898 xenopus tro
158	15	44.1	214	2	06OGR7_CAEBR	06ogr7 caenorhabdi	231	15	44.1	357	2	06LIP0_DROME	06lip0 dirosophila
159	15	44.1	215	2	04QSI8_9LECA	04qsi8 pertusaria	232	15	44.1	357	2	09XV87_CAEBL	09xv87 caenorhabdi
160	15	44.1	217	2	07QY29_GIALA	07qy29 giardia lam	233	15	44.1	362	1	MESD1_HUMAN	09h16 homo sapien
161	15	44.1	218	1	CKSP2_RAT	09j15 rattus norv	234	15	44.1	362	1	0510F9_HUMAN	0510f9 homo sapien
162	15	44.1	220	2	07IT10_CHICK	07it10 gallus gall	235	15	44.1	362	2	07QJL0_ANOCA	07qjl0 anopheles g
163	15	44.1	224	2	05KCO4_CRYNE	05kco4 cryptococcu	236	15	44.1	362	2	04V3U8_DROME	04v3u8 dirosophila
164	15	44.1	225	2	04QSH8_9LECA	04qsh8 pertusaria	237	15	44.1	362	2	05B3V3_RAT	05b3v3 rattus norv
165	15	44.1	225	2	055XU2_CRYNE	055xu2 cryptococcu	238	15	44.1	362	2	0542D0_MOUSE	0542d0 mus musculu
166	15	44.1	226	1	C553_PARDE	P29967 paracoccus	239	15	44.1	371	2	08CFL0_MOUSE	08cfl0 mus musculu
167	15	44.1	226	1	051672_PARDE	051672 paracoccus	240	15	44.1	371	2	0510F9_HUMAN	0510f9 homo sapien
168	15	44.1	227	2	04QSG3_9LECA	04qsg3 pertusaria	241	15	44.1	379	2	09XIV6_ORYSA	09xiv6 oryza sativ
169	15	44.1	227	2	04QSG10_9LECA	04qsg10 pertusaria	242	15	44.1	385	2	07SEFP_ASHCO	07sefp ashyia goss
170	15	44.1	227	2	04QSM6_9ASCO	04qsm6 coccoltrema	243	15	44.1	386	2	07TWC3_TOXCO	07twc3 toxoplasm
171	15	44.1	227	2	04QSU8_9LECA	04qsu8 pertusaria	244	15	44.1	387	2	09VEY9_DROME	09vey9 dirosophila
172	15	44.1	227	2	04QSG8_9LECA	04qsg8 pertusaria	245	15	44.1	387	2	08KQ10_BURCE	08kq10 burkholderi
173	15	44.1	227	2	04QSF6_9LECA	04qsf6 pertusaria	246	15	44.1	387	2	08KQJ7_BURCE	08kqj7 burkholderi
174	15	44.1	227	2	04QSG0_9LECA	04qsg0 pertusaria	247	15	44.1	398	2	09B9Y6_OPHPO	09b9y6 ophiochloa p
175	15	44.1	227	2	04QSI6_9LECA	04qsi6 pertusaria	248	15	44.1	398	2	09B9Y2_9GOBI	09b9y2 acanthogobi
176	15	44.1	227	2	05UTF7_HUMAN	05utef7 homo sapien	249	15	44.1	398	2	09B6C2_LATYA	09b6c2 lateolatrax
177	15	44.1	227	2	09UT7A3_DROME	09ut7a3 dirosophila	250	15	44.1	398	2	09BA00_9GOBI	09ba00 bostrychnus

251	15	44.1	407	2	Q8SGB0	MICSA	Q8sgb0	micropterus	324	15	44.1	688	2	Q4PGR8	USTWA	Q4pgr8	uttilago ma
252	15	44.1	409	2	Q9SUA2	ARATH	Q9sua2	arabidopsis	325	15	44.1	694	2	Q4T2I2	TEITNG	Q4t2i2	tetradon n
253	15	44.1	410	2	Q5HWY7	HUMAN	Q5hwY7	homo sapien	326	15	44.1	695	2	Q7XG03	ORYSA	Q7xg03	oryza sativ
254	15	44.1	411	1	CPRL	YARLI	P45815	yarrowia li	327	15	44.1	695	2	Q94HU6	ORYSA	Q94hu6	oryza sativ
255	15	44.1	415	2	Q5TFE2	ANOGA	Q5tfp2	anopheles g	328	15	44.1	705	2	Q7S136	NEUCR	Q7s136	neurospora
256	15	44.1	416	2	Q6RKQ7	GIBBEZ	Q6rkq7	gibberella	329	15	44.1	705	2	Q81BW1	GITALA	Q81bw1	giardia lam
257	15	44.1	419	2	Q5SRJ3	CRYNE	Q5srj3	cryptococcu	330	15	44.1	713	1	LRPI0	HUMAN	LRPI0	homo sapien
258	15	44.1	420	1	Y461	SYNY3	Q55173	synecocyst	331	15	44.1	713	1	Q5RDI3	PONPY	Q5rdi3	pongo pygma
259	15	44.1	422	2	Q7RO75	GITALA	Q7ro75	giardia lam	332	15	44.1	716	2	Q9UB4	9CILL	Q9ub4	parametium
260	15	44.1	423	2	Q9C8J4	ARATH	Q9c8j4	arabidopsis	333	15	44.1	727	2	Q8VLQ5	ALTSC	Q8vllq5	altomonas
261	15	44.1	428	2	Q4SEM6	TEITNG	Q4sem6	tetradon n	334	15	44.1	731	2	Q8RTZ9	GCANM	Q8rtz9	gcanm
262	15	44.1	432	2	Q9VMD4	DROME	Q9vmd4	drosophila	335	15	44.1	743	2	Q4T6E2	TEITNG	Q4t6e2	teitng
263	15	44.1	443	2	Q5B5Q6	EMENT	Q5b5q6	aspergillus	336	15	44.1	746	2	Q4SPR7	TEITNG	Q4spr7	teitng
264	15	44.1	443	2	Q8LFZ9	ARATH	Q8lfz9	arabidopsis	337	15	44.1	754	2	Q4SSB1	TEITNG	Q4ssb1	teitng
265	15	44.1	443	2	Q93Z67	ARATH	Q93z67	arabidopsis	338	15	44.1	771	2	Q8K4I3	MOUSE	Q8k4i3	mouse
266	15	44.1	459	2	Q9XKP6	CABEL	Q9xkp6	caenorhabdi	339	15	44.1	772	2	Q4WK08	ASPFU	Q4wk08	aspeffu
267	15	44.1	459	2	Q4FZ07	LEITMA	Q4fz07	leishmania	340	15	44.1	799	2	Q5QNF9	ORYSA	Q5qnf9	oryza sativ
268	15	44.1	460	2	Q725N8	DESVI	Q725n8	desulfovibr	341	15	44.1	811	2	Q915B2	HPBYO	Q915b2	hpbyo
269	15	44.1	461	2	Q58317	GYRYP	Q58317	trypanosoma	342	15	44.1	813	2	Q9BLJ2	CIOIN	Q9blj2	ciona intes
270	15	44.1	464	2	Q7R6B2	GITALA	Q7r6b2	giardia lam	343	15	44.1	833	2	Q6J288	ACACA	Q6j288	acanthamoeb
271	15	44.1	465	2	Q4RETO	TEITNG	Q4reto	tetradon n	344	15	44.1	840	1	RNF19	MOUSE	RNF19	mouse
272	15	44.1	470	2	Q8X0Z2	GIBBU	Q8x0z2	gibberella	345	15	44.1	843	2	Q915C3	HPBYO	Q915c3	hpbyo
273	15	44.1	471	2	Q9VMG7	DROME	Q9vmg7	drosophila	346	15	44.1	843	2	Q91572	HPBYO	Q91572	hpbyo
274	15	44.1	485	2	Q5U795	AZOLI	Q5u795	azospirillum	347	15	44.1	843	2	Q81141	HPBYO	Q81141	hpbyo
275	15	44.1	489	1	ZIFI	CABEL	P34482	caenorhabdi	348	15	44.1	843	2	Q8JXG0	HPBYO	Q8jxg0	hpbyo
276	15	44.1	497	2	Q7UPM3	RHOBA	Q7upm3	rhodospirell	349	15	44.1	869	2	Q8FGB3	ECOL6	Q8fgb3	escherichia
277	15	44.1	500	2	Q7JHJ6	RHOPA	Q7jhj6	rhodospirell	350	15	44.1	869	2	Q8PMF6	XANAC	Q8pmf6	xanthomonas
278	15	44.1	503	2	Q4YZM1	PLABE	Q4yzm1	plasmidium	351	15	44.1	885	2	Q6B3B9	SPROT	Q6b3b9	spiroplasma
279	15	44.1	505	2	Q4FKT0	GYRYP	Q4fkt0	trypanosoma	352	15	44.1	888	2	Q81PZ1	DROME	Q81pz1	drome
280	15	44.1	516	2	Q54C18	DICDI	Q54c18	dictyostell	353	15	44.1	904	2	Q4QAJ0	LEITMA	Q4qaj0	leishmania
281	15	44.1	516	2	Q8P76	VIBPA	Q8p76	vibrio para	354	15	44.1	935	2	Q94719	PARTT	Q94719	partecium
282	15	44.1	520	2	Q8N1N4	HUMAN	Q8n1n4	homo sapien	355	15	44.1	937	2	Q5N1I1	FRATF	Q5n1i1	fratf
283	15	44.1	521	2	Q7RTT2	HUMAN	Q7rtc2	homo sapien	356	15	44.1	943	2	Q60DG4	ORYSA	Q60dg4	oryza sativ
284	15	44.1	522	2	Q5BAN2	EMENT	Q5ban2	aspergillus	357	15	44.1	948	2	Q5NKH5	AZORR	Q5nkh5	azotarcus sp
285	15	44.1	522	2	Q55ZQ1	CRYNE	Q55zq1	cryptococcu	358	15	44.1	969	2	Q8BR56	MOUSE	Q8br56	mouse
286	15	44.1	525	2	Q5KPI3	CRYNE	Q5kpi3	cryptococcu	359	15	44.1	984	2	Q81PZ2	DROME	Q81pz2	drome
287	15	44.1	527	2	Q6ETD6	ORYSA	Q6etd6	oryza sativ	360	15	44.1	988	2	Q7QYU1	GITALA	Q7qyu1	giardia lam
288	15	44.1	531	1	PGRL1	RALSO	P58598	raistrontia s	361	15	44.1	993	2	Q9VOL8	DROME	Q9vol8	drome
289	15	44.1	538	2	Q5IRL7	MAGGR	Q5irl7	magnaporthe	362	15	44.1	1002	1	SPHR	AMEPV	SPHR	amepv
290	15	44.1	541	2	Q4WJAO	ASPFU	Q4wja0	aspergillus	363	15	44.1	1003	2	Q5AMZ7	DICDI	Q5amz7	dictyostell
291	15	44.1	544	2	Q941V2	ORYSA	Q941v2	oryza sativ	364	15	44.1	1013	2	Q4SRB6	TEITNG	Q4srb6	teitng
292	15	44.1	549	2	P77818	AGRYVI	P77818	halocynthia	365	15	44.1	1023	2	Q5JUB6	HUMAN	Q5jub6	homo sapien
293	15	44.1	563	2	Q6V9W0	RAT	Q6v9w0	rat	366	15	44.1	1025	2	Q7R6U7	GITALA	Q7r6u7	giardia lam
294	15	44.1	564	2	Q9YQZ0	9HERP	Q9yqz0	ratid herpe	367	15	44.1	1025	2	Q94EN9	CHLRE	Q94en9	chlamydomon
295	15	44.1	565	1	SRK23	MOUSE	Q940g2	mus muscullu	368	15	44.1	1039	2	Q610W2	CABER	Q610w2	caber
296	15	44.1	569	2	Q9CBB6	MYCLE	Q9cbb6	mycobacteri	369	15	44.1	1043	2	Q17644	CABEL	Q17644	caenorhabdi
297	15	44.1	574	2	Q8MXM1	HALRO	Q8mxm1	halocynthia	370	15	44.1	1057	2	Q4N4P8	THEPA	Q4n4p8	theilertia p
298	15	44.1	587	2	Q9VHJ6	DROME	Q9vhj6	drosophila	371	15	44.1	1061	2	Q4XQJ4	PLACH	Q4xqj4	plasmidium
299	15	44.1	589	2	Q4P936	USTWA	Q4p936	uttilago ma	372	15	44.1	1063	2	Q7QJ010	GITALA	Q7qj010	giardia lam
300	15	44.1	593	2	Q818V8	GITALA	Q818v8	giardia lam	373	15	44.1	1068	2	Q4Q8B6	LEITMA	Q4q8b6	leishmania
301	15	44.1	593	2	Q81MC6	ORYSA	Q81mc6	oryza sativ	374	15	44.1	1070	2	Q7R2W4	GITALA	Q7r2w4	giardia lam
302	15	44.1	595	2	Q5VXL7	HUMAN	Q5vxl7	homo sapien	375	15	44.1	1089	2	Q4Q9C2	LEITMA	Q4q9c2	leishmania
303	15	44.1	599	2	Q8W4D7	ARATH	Q8w4d7	arabidopsis	376	15	44.1	1103	2	Q8MR15	DROME	Q8mr15	drome
304	15	44.1	602	2	Q6MG12	NEUCR	Q6mg12	neurospora	377	15	44.1	1146	2	Q7OXI5	GITALA	Q7oxi5	giardia lam
305	15	44.1	605	2	Q9SFU8	ARATH	Q9sfu8	arabidopsis	378	15	44.1	1173	2	Q5U4V9	XENLA	Q5u4v9	xenopus lae
306	15	44.1	608	2	Q7SDR2	NEUCR	Q7sdr2	neurospora	379	15	44.1	1186	1	C14A	BACTS	C14A	bacillus th
307	15	44.1	609	2	Q64BX6	9ARCH	Q64bx6	uncultured	380	15	44.1	1189	2	Q4NUU7	9DELT	Q4nuu7	anaeromyxob
308	15	44.1	609	2	Q64DG6	9ARCH	Q64dg6	uncultured	381	15	44.1	1191	2	Q8RL60	PSEFL	Q8rl60	pusillomonas
309	15	44.1	614	2	Q8GZ21	ARATH	Q8gz21	arabidopsis	382	15	44.1	1216	2	Q4PDB3	USTWA	Q4pdb3	uttilago ma
310	15	44.1	622	2	Q8C9V4	MOUSE	Q8c9v4	mus muscullu	383	15	44.1	1232	2	Q4T1X3	TEITNG	Q4t1x3	teitng
311	15	44.1	633	2	Q818W0	GITALA	Q818w0	giardia lam	384	15	44.1	1236	2	Q8OTKO	MOUSE	Q8otko	mus muscullu
312	15	44.1	635	2	Q7QPO7	GITALA	Q7qp07	giardia lam	385	15	44.1	1249	2	Q57VB8	GYRYP	Q57vb8	trypanosoma
313	15	44.1	637	2	Q4UIV3	THEAN	Q4uiv3	theilertia a	386	15	44.1	1253	2	Q4TOS1	TEITNG	Q4tos1	tetradon n
314	15	44.1	638	2	Q7QOC4	GITALA	Q7qoc4	giardia lam	387	15	44.1	1285	2	Q8DA47	VIBVU	Q8da47	vibrio vuln
315	15	44.1	639	2	Q8PMN5	METMA	Q8pmn5	methanosarc	388	15	44.1	1297	2	Q7MKJ2	VIBVU	Q7mkj2	vibrio vuln
316	15	44.1	639	2	Q8TTZ6	METMA	Q8ttz6	methanosarc	389	15	44.1	1324	2	Q8SYK2	DROME	Q8syk2	drome
317	15	44.1	640	2	Q8C121	MOUSE	Q8c121	mus muscullu	390	15	44.1	1324	2	Q9VTR8	DROME	Q9vtr8	drome
318	15	44.1	644	2	Q7QWT5	GITALA	Q7qwt5	giardia lam	391	15	44.1	1343	2	Q4IK63	GIBBEZ	Q4ik63	gibberella
319	15	44.1	659	2	Q6CGW2	TARLI	Q6cgw2	yarrowia li	392	15	44.1	1353	2	Q6R2R5	ARATH	Q6r2r5	arabidopsis
320	15	44.1	659	2	Q4N6R1	THEPA	Q4n6r1	theilertia p	393	15	44.1	1361	2	Q9NGV2	DROME	Q9ngv2	drosophila
321	15	44.1	669	2	Q4WMZ9	ASPFU	Q4wmz9	aspergillus	394	15	44.1	1361	2	Q9V714	DROME	Q9v714	drome
322	15	44.1	679	2	Q6PUC7	ANOGA	Q6puc7	anopheles g	395	15	44.1	1364	2	Q9SW45	ARATH	Q9sw45	arabidopsis
323	15	44.1	688	2	Q96VF2	USTWA	Q96vf2	uttilago ma	396	15	44.1	1376	2	Q8X1P2	PODAN	Q8x1p2	podospira a

397	15	44.1	1414	2	Q40E45_LEIMA	Q4qE45 leishmania	470	15	44.1	10495	2	Q4RE92_TETNG	Q4re92 tetraodon n
398	15	44.1	1429	2	Q7PW3_ASPTU	Q7pw3 anopheles g	471	15	44.1	34350	2	Q8WZ42_HUMAN	Q8wz42 homo sapien
399	15	44.1	1447	2	Q4WPJ3_ASPFU	Q4wpj3 aspergillus	472	14	41.2	21	2	Q7M276_GASCO	Q7m276 gastroclon
400	15	44.1	1451	1	EM30_ARATH	Q42510 arabidopsis	473	14	41.2	29	2	Q52D63_MAGGR	Q52d63 magnaporthe
401	15	44.1	1468	2	Q80TF6_MOUSE	Q80tf6 mus musculu	474	14	41.2	29	2	Q4TIR3_TETNG	Q4tir3 tetraodon n
402	15	44.1	1514	2	Q4FX65_LEIMA	Q4fx65 leishmania	475	14	41.2	44	2	Q5TUB2_ANOGA	Q5tub2 anopheles g
403	15	44.1	1526	2	Q7RLI5_PLAYO	Q7rli5 plasmodium	476	14	41.2	46	2	Q7RXU9_NEUCR	Q7rxu9 neotropo
404	15	44.1	1533	2	Q7RIL0_PLAYO	Q7ril0 plasmodium	477	14	41.2	52	2	Q7UFK9_RHOBA	Q7ufk9 rhodopirell
405	15	44.1	1553	2	Q4Q5Q2_LEIMA	Q4q5q2 leishmania	478	14	41.2	53	2	Q7TSS9_RAT	Q7tss9 rattus norv
406	15	44.1	1599	2	Q4Q5Q2_LEIMA	Q4q5q2 leishmania	479	14	41.2	55	2	Q5VIZ4_AERHY	Q5viz4 aeromonas h
407	15	44.1	1624	2	Q7RM69_PLAYO	Q7rm69 plasmodium	480	14	41.2	57	2	Q4TIC2_TETNG	Q4tic2 tetraodon n
408	15	44.1	1625	2	Q4T364_TETNG	Q4t364 tetraodon n	481	14	41.2	58	2	Q5EY67_ADBOT	Q5ey67 human adeno
409	15	44.1	1651	2	Q7RNX5_PLAYO	Q7rnx5 plasmodium	482	14	41.2	64	2	Q4YVW8_PLASE	Q4yvW8 plasmodium
410	15	44.1	1722	2	Q4N0X1_9DELT	Q4n0x1 anaeromyxob	483	14	41.2	68	2	Q4YBP5_PLASE	Q4ybp5 plasmodium
411	15	44.1	1729	2	Q7XGP6_ORYSA	Q7xgp6 oryza sativ	484	14	41.2	70	2	Q84Z47_ORYSA	Q84z47 oryza sativ
412	15	44.1	1729	2	Q8LZT0_ORYSA	Q8lzt0 oryza sativ	485	14	41.2	73	2	Q8RVIL_PIPNS	Q8rvil pinus pinas
413	15	44.1	1755	2	Q6RKT6_BOTCI	Q6rkt6 botrytis ci	486	14	41.2	74	2	Q87AK8_XYLFT	Q87ak8 xylecta fas
414	15	44.1	1782	2	Q4WQZ5_ASPFU	Q4wqz5 aspergillus	487	14	41.2	76	2	Q5TYD0_ANOGA	Q5tyd0 anopheles g
415	15	44.1	1782	2	Q6X012_SOLIN	Q6x012 solenopais	488	14	41.2	76	2	Q89W35_BRALU	Q89w35 bradyrhizob
416	15	44.1	1793	2	Q4FMM9_LEIMA	Q4fmm9 leishmania	489	14	41.2	77	2	Q9T023_GCAUD	Q9t023 bacterioph
417	15	44.1	1806	2	Q5BH30_LEIMA	Q5bh30 aspergillus	490	14	41.2	78	2	Q7UHI8_RHOBA	Q7uhi8 rhodopirell
418	15	44.1	1893	2	Q4U0S1_BRARE	Q4u0s1 brachydanio	491	14	41.2	80	2	Q67N17_SYMTH	Q67n17 symbiobacte
419	15	44.1	1944	2	Q6WS81_9ACTO	Q6ws81 actinomadar	492	14	41.2	80	2	Q4T6X9_TETNG	Q4t6x9 tetraodon n
420	15	44.1	1946	2	Q54UJ3_DICDI	Q54j3 dictyosteli	493	14	41.2	82	2	Q4SIR2_TETNG	Q4sir2 tetraodon n
421	15	44.1	1958	2	Q84HN4_9ACTO	Q84hn4 klataetospo	494	14	41.2	83	2	Q65MR0_ORYSA	Q65mr0 oryza sativ
422	15	44.1	1961	2	Q84HM9_9ACTO	Q84hm9 streptomyce	495	14	41.2	86	2	Q5P1S6_AZOSE	Q5p1s6 azoarcus sp
423	15	44.1	1976	2	Q7PUW1_ANOGA	Q7puw1 anopheles g	496	14	41.2	88	2	P97023_BRELN	P97023 brevibacter
424	15	44.1	1977	2	Q8MU72_ANOGA	Q8mu72 anopheles g	497	14	41.2	90	2	Q9PFI7_XYLPA	Q9pfi7 xylecta fas
425	15	44.1	1978	2	Q8T5H2_ANOGA	Q8t5h2 anopheles g	498	14	41.2	90	2	Q57D56_BRUBA	Q57d56 bruceella ab
426	15	44.1	1988	2	Q6RK18_BOTCI	Q6rk18 botrytis ci	499	14	41.2	91	2	Q4XC49_PLACH	Q4xc49 plasmodium
427	15	44.1	1996	2	Q4T528_TETNG	Q4t528 tetraodon n	500	14	41.2	91	2	Q91667_XENIA	Q91667 xenopus lae
428	15	44.1	2043	2	Q4Q510_LEIMA	Q4q510 leishmania	501	14	41.2	93	1	RS16_CANTP	RS16 candidatus
429	15	44.1	2087	2	Q4S488_TETNG	Q4s488 tetraodon n	502	14	41.2	93	2	Q872R7_NEUCR	Q872r7 neotropo
430	15	44.1	2098	2	P91443_CAEBL	P91443 caenorhabdi	503	14	41.2	93	2	Q6Z6F3_ORYSA	Q6z6f3 oryza sativ
431	15	44.1	2099	2	Q622K8_CAEBR	Q622k8 caenorhabdi	504	14	41.2	94	2	P97PC59_MYCSM	P97pc59 mycobacteri
432	15	44.1	2106	2	Q6RWD9_NECHA	Q6rwd9 neectria hae	505	14	41.2	97	2	P91216_CAEBL	P91216 caenorhabdi
433	15	44.1	2112	2	Q8WPL0_9UROC	Q8wpl0 oikopleura	506	14	41.2	97	2	Q53CUI_9GAMA	Q53cui macaca fusc
434	15	44.1	2117	2	Q71449_GEOSL	Q71449 geobacter s	507	14	41.2	98	2	Q7S7T5_NEUCR	Q7s7t5 neotropo
435	15	44.1	2159	2	Q9Y8A6_9PEZI	Q9y8a6 nodulispoti	508	14	41.2	100	2	Q9YFR0_AERPE	Q9yfr0 aeteryptus p
436	15	44.1	2159	2	Q8GIU5_DICDI	Q8giu5 dictyosteli	509	14	41.2	101	2	Q9DXF8_9CIRC	Q9dxF8 beak and fe
437	15	44.1	2162	2	Q5S1N7_DICDI	Q5s1n7 dictyosteli	510	14	41.2	101	2	Q9DXG4_9CIRC	Q9dxG4 beak and fe
438	15	44.1	2162	2	Q8NK46_9PEZI	Q8nk46 xylaria sp.	511	14	41.2	101	2	Q9DXG7_9CIRC	Q9dxG7 beak and fe
439	15	44.1	2162	2	Q9VQW0_DROME	Q9vqwu dtrosophila	512	14	41.2	102	2	Q9MRK6_9RALV	Q9mrk6 chlamydomophi
440	15	44.1	2173	2	Q51PP8_MAGGR	Q51pp8 magnaporthe	513	14	41.2	102	2	Q824W5_CHLGY	Q824w5 chlamydophi
441	15	44.1	2176	2	Q46112_DROME	Q46112 dtrosophila	514	14	41.2	103	1	Y040_MYCPN	Y05074 mycoplasma
442	15	44.1	2181	1	STCA_EMENTI	Q12397 emeritella	515	14	41.2	103	2	Q5YF63_VIVRU	Q5yF63 rock bream
443	15	44.1	2187	2	P79068_GIOLA	P79068 glomerella	516	14	41.2	104	2	Q13588_YEAGT	Q13588 saccharomyc
444	15	44.1	2187	2	Q8TGD7_ASPTU	Q8tgd7 aspergillus	517	14	41.2	105	2	Q4TPW2_9SPPN	Q4tpw2 erythrobact
445	15	44.1	2188	2	Q6XR12_9PEZI	Q6xr12 ceratocysti	518	14	41.2	107	1	YB56_YEAST	P8309 saccharomyc
446	15	44.1	2203	2	Q7S736_NEUCR	Q7s736 neotropo	519	14	41.2	107	2	Q05413_YEAGT	Q05413 saccharomyc
447	15	44.1	2238	2	Q5AV55_EMENTI	Q5av55 aspergillus	520	14	41.2	108	2	Q07521_YEAGT	Q07521 saccharomyc
448	15	44.1	2275	2	Q6RKE4_COCHC	Q6rke4 cochllopolu	521	14	41.2	109	2	Q6K452_ORYSA	Q6k452 oryza sativ
449	15	44.1	2397	2	Q6Q7Y4_PARPR	Q6q7y4 paramectium	522	14	41.2	110	2	Q82YH0_STRAW	Q82yh0 streptomyc
450	15	44.1	2511	2	Q4NXX5_9DELT	Q4nxx5 anaeromyxob	523	14	41.2	111	2	Q7SC86_NEUCR	Q7sc86 neotropo
451	15	44.1	2517	2	Q5A232_EMENTI	Q5a232 aspergillus	524	14	41.2	111	2	Q8IX35_HUMAN	Q8ix35 homo sapien
452	15	44.1	2533	2	P90589_PARTE	P90589 paramectium	525	14	41.2	112	2	P72474_STRMU	P72474 streptococc
453	15	44.1	2533	2	Q27183_PARTE	Q27183 paramectium	526	14	41.2	112	2	Q51906_DESVU	Q51906 desulfovibr
454	15	44.1	2543	2	P90649_PARPR	P90649 paramectium	527	14	41.2	113	2	Q4V572_DROME	Q4v572 dtrosophila
455	15	44.1	2556	2	Q7QUR3_ANOGA	Q7qj93 anopheles g	528	14	41.2	113	2	Q4V572_DROME	Q4v572 dtrosophila
456	15	44.1	2644	2	Q6RK12_BOTCI	Q6rk12 botrytis ci	529	14	41.2	113	2	Q56754_9HEPC	Q56754 pinus radia
457	15	44.1	2703	1	NOTCH_DROME	P07207 dtrosophila	530	14	41.2	115	2	Q97682_SULTO	Q97682 bullitobus
458	15	44.1	2717	2	Q94710_PARTE	Q94710 paramectium	531	14	41.2	116	2	Q8M633_9CAUD	Q8m633 bacterioph
459	15	44.1	2729	2	Q6PQK6_PARTE	Q6pqk6 paramectium	532	14	41.2	118	2	Q868W8_PLAPA	Q868w8 plasmodium
460	15	44.1	2793	2	Q59IS5_BRARE	Q59ijs brachydanio	533	14	41.2	118	2	Q8RFE3_FUSNN	Q8rfE3 fusobacteri
461	15	44.1	2836	2	Q5BRE6_EMENTI	Q5bre6 aspergillus	534	14	41.2	119	2	Q4OAH0_LEISMA	Q4oah0 leishmania
462	15	44.1	2836	2	Q9VSU8_DROME	Q9vsu8 dtrosophila	535	14	41.2	120	2	Q51901_DESEA	Q51901 desulfovibr
463	15	44.1	4039	2	Q4KCD6_PSEF5	Q4kcd6 pseudomonas	536	14	41.2	120	2	Q7PEK3_FUSNV	Q7pek3 fusobacteri
464	15	44.1	4376	2	Q5MP07_9BACT	Q5mp07 symbioton ba	537	14	41.2	121	2	Q84ZU5_ORYSA	Q84zU5 oryza sativ
465	15	44.1	4767	2	Q17301_CAEBR	Q17301 caenorhabdi	538	14	41.2	121	2	Q4SZW9_TETNG	Q4szw9 tetraodon n
466	15	44.1	4881	2	Q6TAR6_9BACT	Q6tar6 symbioton ba	539	14	41.2	122	2	Q69HS6_CIOIN	Q69hs6 ctiona intes
467	15	44.1	7743	2	Q618F4_CAEBR	Q618f4 caenorhabdi	540	14	41.2	123	2	Q6Y9W1_NICSY	Q6y9w1 nicotiana s
468	15	44.1	7829	2	Q18559_CAEBL	Q18559 caenorhabdi	541	14	41.2	123	2	Q51904_9DELT	Q51904 desulfovibr
469	15	44.1	7962	2	Q10465_HUMAN	Q10465 homo sapien	542	14	41.2	123	2	Q8VKK8_MYCTU	Q8vkk8 mycobacteri

543	14	41.2	123	2	093464_CARAU	093464	catarsius a	616	14	41.2	158	2	0641h1_9CIRC	0641h1	beak and fe
544	14	41.2	124	2	07S9A8_NEUCR	07S9A8	neuropora	617	14	41.2	158	2	09DXF2_9CIRC	09DXF2	beak and fe
545	14	41.2	124	2	08NB52_HUMAN	08NB52	homo sapien	618	14	41.2	158	2	09DXF5_9CIRC	09DXF5	beak and fe
546	14	41.2	124	2	051916_9BELT	051916	unidentifie	619	14	41.2	159	2	06RZ20_9PEZT	06RZ20	periconia m
547	14	41.2	124	2	04V2P3_BURMA	04V2P3	burkholderi	620	14	41.2	159	2	09PDM1_XYTRA	09PDM1	xyella fas
548	14	41.2	124	2	06IWP5_9EURO	06IWP5	petromyces	621	14	41.2	160	2	04SYU6_TETNG	04SYU6	tetradon n
549	14	41.2	125	2	06E1W8_9EURO	06E1W8	aspergillus	622	14	41.2	161	2	06ZU72_HUMAN	06ZU72	homo sapien
550	14	41.2	125	2	06E1W9_9EURO	06E1W9	petromyces	623	14	41.2	161	2	09SL78_MACMU	09SL78	macaca mula
551	14	41.2	125	2	06E1X0_9EURO	06E1X0	petromyces	624	14	41.2	162	2	09IGT6_ADEP3	09IGT6	porcine ade
552	14	41.2	125	2	07R1Y6_GITALA	07R1Y6	giardia lam	625	14	41.2	162	2	04TH15_TETNG	04TH15	tetradon n
553	14	41.2	126	2	09JH18_MOUSE	09JH18	mus musculu	626	14	41.2	163	2	06RZ18_9EURO	06RZ18	telatromyces
554	14	41.2	129	1	LYSC1_CANFA	P81708	canis famli	627	14	41.2	163	2	09P2S0_HUMAN	09P2S0	homo sapien
555	14	41.2	129	1	05C4S6_SCHJA	05C4S6	schistosoma	628	14	41.2	163	2	09P7R9_GITALA	09P7R9	giardia lam
556	14	41.2	130	2	06RZ24_9PLEBO	06RZ24	stremphylium	629	14	41.2	163	2	06XK90_ORYSA	06XK90	oryza sativ
557	14	41.2	131	1	NEU22_4NSAN	P169230	anser anser	630	14	41.2	164	2	06TQ01_9LECA	06TQ01	lecanora ma
558	14	41.2	132	2	07S107_NEUCR	07S107	neuropora	631	14	41.2	164	2	06TQ00_9LECA	06TQ00	lecanora na
559	14	41.2	132	2	09VL37_DROME	09VL37	dirosophila	632	14	41.2	164	2	06TQ02_9LECA	06TQ02	lecanora fi
560	14	41.2	132	2	08G4M6_BIFLO	08G4M6	bifidobacte	633	14	41.2	164	2	06TQ03_9PEZT	06TQ03	dermatocarp
561	14	41.2	133	1	YB9U_YEAST	P38350	saccharomyc	634	14	41.2	164	2	0445B6_CABEL	0445B6	caenorhabdi
562	14	41.2	133	2	0870B4_9FRUNG	0870B4	plyomyces s	635	14	41.2	165	2	05SNE2_ORYSA	05SNE2	oryza sativ
563	14	41.2	133	2	05TVJ5_4NOGA	05TVJ5	anopheles g	636	14	41.2	165	2	0445B7_CABEL	0445B7	caenorhabdi
564	14	41.2	134	2	06TV1_HUMAN	06TV1	homo sapien	637	14	41.2	166	2	09BHX2_9ANNE	09BHX2	sabella spa
565	14	41.2	134	2	06ZFU3_ORYSA	06ZFU3	oryza sativ	638	14	41.2	167	2	05H5C2_XANOR	05H5C2	xanthomonas
566	14	41.2	135	2	05Q994_IXOSC	05Q994	ixodes scap	639	14	41.2	167	2	08VK95_MYCTU	08VK95	mycobacteri
567	14	41.2	135	2	018158_CABEL	018158	caenorhabdi	640	14	41.2	168	2	08H2U1_ORYSA	08H2U1	oryza sativ
568	14	41.2	135	2	092PT0_RHIME	092PT0	rhizobium m	641	14	41.2	169	2	08H0A4_ORYSA	08H0A4	oryza sativ
569	14	41.2	136	2	061LL3_DROME	061LL3	dirosophila	642	14	41.2	169	2	06REU4_9FLOR	06REU4	halymenia s
570	14	41.2	136	2	08TUJ9_METAC	08TUJ9	methanosc	643	14	41.2	169	2	08ZG98_STRAW	08ZG98	strepomyce
571	14	41.2	137	2	07Q4X7_4NOGA	07Q4X7	anopheles g	644	14	41.2	170	2	06ZWC5_HUMAN	06ZWC5	homo sapien
572	14	41.2	137	2	08BWD7_MOUSE	08BWD7	m mus muscu	645	14	41.2	170	2	07UZZ4_RHOBA	07UZZ4	rhodopirell
573	14	41.2	139	2	06RZ21_9EURO	06RZ21	dichotomomy	646	14	41.2	171	2	08LNA4_ORYSA	08LNA4	oryza sativ
574	14	41.2	139	2	000830_LEIMA	000830	leishmania	647	14	41.2	171	2	09LBO4_9ACTO	09LBO4	actreptomyce
575	14	41.2	141	2	077076_HAECC	077076	haemochus	648	14	41.2	172	2	08SZV2_DROME	08SZV2	dirosophila
576	14	41.2	141	2	084MP8_ORYSA	084MP8	oryza sativ	649	14	41.2	174	2	05B0Z5_AEMET	05B0Z5	aspergillus
577	14	41.2	143	1	PONA_DICDI	P54660	dictyosteli	650	14	41.2	174	2	06RZ25_9PLEBO	06RZ25	westernydel
578	14	41.2	143	2	0585A5_9TRYRP	0585A5	trypanosoma	651	14	41.2	174	2	081PD4_DROME	081PD4	dirosophila
579	14	41.2	143	2	054BG4_DICDI	054BG4	dictyosteli	652	14	41.2	175	1	SNAC_STRPR	SNAC	strepomyce
580	14	41.2	143	2	06NA71_RHOPA	06NA71	rhodopseudo	653	14	41.2	175	2	05STU7_CRYNE	05STU7	cryptococcu
581	14	41.2	145	2	07R3S4_GITALA	07R3S4	giardia lam	654	14	41.2	175	2	08GN71_BIFLO	08GN71	bifidobacte
582	14	41.2	148	2	052CU3_MAGGR	052CU3	magnaporthe	655	14	41.2	175	2	05H2G2_XANOR	05H2G2	xanthomonas
583	14	41.2	148	2	08EV25_MYCPE	08EV25	mycoplasma	656	14	41.2	175	2	07WX94_4LCEU	07WX94	atcaligenes
584	14	41.2	149	2	06RZ22_9PEZT	06RZ22	cladosporiu	657	14	41.2	175	2	082PB6_STRAW	082PB6	strepomyce
585	14	41.2	149	2	07OJ01_4NOGA	07OJ01	anopheles g	658	14	41.2	176	2	06RZ19_9ASCO	06RZ19	pleistophoti
586	14	41.2	149	2	07XG18_ORYSA	07XG18	oryza sativ	659	14	41.2	177	1	PHEB_POLBO	PHEB	polypora pu
587	14	41.2	149	2	094102_ORYSA	094102	oryza sativ	660	14	41.2	177	1	PHEB_PORPU	PHEB	porphyra te
588	14	41.2	149	2	08AV79_BRARE	08AV79	brachydanio	661	14	41.2	177	1	PHEB_PORPE	PHEB	porphyra ye
589	14	41.2	150	2	06RZ27_9EURO	06RZ27	penicillillium	662	14	41.2	177	1	09YD43_AERPE	09YD43	aeropyrum p
590	14	41.2	150	2	06TGP9_9LECA	06TGP9	lecanora st	663	14	41.2	177	2	04WU6_4SPPU	04WU6	aspergillus
591	14	41.2	150	2	061641_ONCOC	061641	onchocerca	664	14	41.2	177	2	08NNA9_HUMAN	08NNA9	homo sapien
592	14	41.2	150	2	05NAT9_ORYSA	05NAT9	oryza sativ	665	14	41.2	177	2	09MRW8_GRALE	09MRW8	gracilaria
593	14	41.2	151	2	07PKW7_4NOGA	07PKW7	anopheles g	666	14	41.2	177	2	07SIP9_GRACH	07SIP9	grach
594	14	41.2	152	2	06RZ15_4SPNG	06RZ15	aspergillus	667	14	41.2	177	2	068FM1_MOUSE	068FM1	mouse
595	14	41.2	154	2	09YA56_AERPE	09YA56	aeropyrum p	668	14	41.2	178	2	09UAV9_CABEL	09UAV9	caenorhabdi
596	14	41.2	155	2	0857R5_9CAUD	0857R5	mycobacteri	669	14	41.2	179	2	060WM3_CABER	060WM3	caenorhabdi
597	14	41.2	156	1	PPDA_ECOLI	P33554	escherichia	670	14	41.2	179	2	067328_9INFA	067328	influenza a
598	14	41.2	156	2	0164Z2_CABEL	0164Z2	caenorhabdi	671	14	41.2	179	2	07YT17_CABEL	07YT17	caenorhabdi
599	14	41.2	156	2	06SEV4_9RACT	06SEV4	uncultured	672	14	41.2	181	2	04LZX2_9BURK	04LZX2	burkholderi
600	14	41.2	156	2	057KX8_SALPA	057KX8	salmonella	673	14	41.2	181	2	08C8W0_MOUSE	08C8W0	mus musculu
601	14	41.2	156	2	05PEM2_4SALPA	05PEM2	salmonella	674	14	41.2	182	2	04TT91_CABEL	04TT91	caenorhabdi
602	14	41.2	156	2	08ZMB0_SALTY	08ZMB0	salmonella	675	14	41.2	183	2	069T53_ORYSA	069T53	oryza sativ
603	14	41.2	156	2	08Z413_SALTY	08Z413	salmonella	676	14	41.2	184	2	08H447_ORYSA	08H447	oryza sativ
604	14	41.2	157	2	025068_HAIR	025068	haematobia	677	14	41.2	184	2	05TWH6_4NOGA	05TWH6	anopheles g
605	14	41.2	157	2	06XHY2_DROYA	06XHY2	dirosophila	678	14	41.2	185	2	05U18_BOVIN	05U18	bov
606	14	41.2	157	2	09BH65_GITALA	09BH65	giardia lam	679	14	41.2	188	1	DUS18_HUMAN	DUS18	homo sapien
607	14	41.2	157	2	09B1L6_GITALA	09B1L6	giardia lam	680	14	41.2	188	1	DUS18_MOUSE	DUS18	mus musculu
608	14	41.2	157	2	06AD65_LEIXX	06AD65	leifsonia x	681	14	41.2	188	1	DUS18_PONPY	DUS18	pungo pygma
609	14	41.2	158	2	006775_HAIR	006775	haematobia	682	14	41.2	188	2	04R3R3_MACFA	04R3R3	macaca fasc
610	14	41.2	158	2	09B1L8_GITALA	09B1L8	giardia lam	683	14	41.2	189	2	058LRF_9CAUD	058LRF	cyanocephale
611	14	41.2	158	2	05SEB2_BACCD	05SEB2	bacillus li	684	14	41.2	189	2	09DPD8_MOUSE	09DPD8	mus musculu
612	14	41.2	158	2	090236_9CIRC	090236	beak and fe	685	14	41.2	189	2	052046_9ZZZZ	052046	plasmid phv
613	14	41.2	158	2	0641FE_9CIRC	0641FE	beak and fe	686	14	41.2	189	2	087TJ5_VIBPA	087TJ5	vibrio para
614	14	41.2	158	2	0641GO_9CIRC	0641GO	beak and fe	687	14	41.2	190	2			
615	14	41.2	158	2	0641G7_9CIRC	0641G7	beak and fe	688	14	41.2	191	2	04MSA1_BACCE	04MSA1	bacillus ce

689	14	41.2	193	2	Q4PD06_USTMA	Q4pd06_ustliago ma	762	14	41.2	221	1	OAZ1_MESAU	P70112_mesocricetu
690	14	41.2	193	2	O5TV08_ANOGA	O5tv08_anophies g	763	14	41.2	221	2	O4Q5J7_9LECA	O4q5j7_pertusaria
691	14	41.2	193	2	O9EM12_9CORO	O9em12_avian infec	764	14	41.2	221	2	O8KENS_CHITE	O8kens_chlorobium
692	14	41.2	195	2	O54199_STRGR	O54199_streptomyce	765	14	41.2	221	2	O5XJN5_BARTE	O5xjn5_brachydario
693	14	41.2	197	2	O4WPH8_ASPFL	O4wfh8_aspergillus	766	14	41.2	222	2	O94722_PARTE	O94722_parmecium
694	14	41.2	197	2	Q7R0U0_GIATA	Q7r0j0_giardia lam	767	14	41.2	222	2	O56PH2_RUEGE	O56ph2_ruegeria ge
695	14	41.2	197	2	O44585_CAEEL	O44585_caenorhabdi	768	14	41.2	224	2	O56PB1_9METZ	O56pb1_uncultured
696	14	41.2	197	2	O8GGS6_STRAZ	O8ggs6_streptomyce	769	14	41.2	225	2	O20590_CAEEL	O20590_caenorhabdi
697	14	41.2	198	1	DUS14_HUMAN	O95147_homo sapien	770	14	41.2	226	2	O5PC56_SALPA	O5pc56_salmonella
698	14	41.2	198	1	DUS14_MOUSE	O91147_homo musculu	771	14	41.2	226	2	O4RDM1_TETNG	O4rdm1_tetradodon n
699	14	41.2	198	2	O6F136_HUMAN	O6f136_homo sapien	772	14	41.2	227	2	O4QSG1_9LECA	O4qsg1_pertusaria
700	14	41.2	198	2	O8VKX4_MYCTU	O8vkx4_mycobacteri	773	14	41.2	227	2	O4QST1_9LECA	O4qst1_pertusaria
701	14	41.2	198	2	O5SV91_MOUSE	O5sv91_mus musculu	774	14	41.2	227	2	O4QSU1_9LECA	O4qsu1_pertusaria
702	14	41.2	198	2	Q7TPY1_MOUSE	Q7tpy1_mus musculu	775	14	41.2	227	2	O4QSN0_9ASCO	O4qsn0_coccotrema
703	14	41.2	199	1	NITI1_YEAST	P40447_saccharomyc	776	14	41.2	227	2	O4QSM9_9ASCO	O4qsm9_coccotrema
704	14	41.2	199	2	O05729_YEAST	O05729_saccharomyc	777	14	41.2	227	2	O4QSM8_9ASCO	O4qsm8_coccotrema
705	14	41.2	199	2	O6O5A1_YEAST	O6o5a1_saccharomyc	778	14	41.2	227	2	O4QSW7_9ASCO	O4qsw7_coccotrema
706	14	41.2	201	2	Q7PUC4_ANOGA	Q7pug4_anopheles g	779	14	41.2	227	2	O4QSM4_9LECA	O4qsm4_ochrolechia
707	14	41.2	201	2	O6DS56_BRARE	O6dgs6_brachydario	780	14	41.2	227	2	O4QSM3_9LECA	O4qsm3_ochrolechia
708	14	41.2	202	2	O4QJ05_9LECA	O4qej5_pertusaria	781	14	41.2	227	2	O4QSL8_9LECA	O4qsl8_ochrolechia
709	14	41.2	202	2	Q4Q3M4_LEIMA	O4q3m4_leishmania	782	14	41.2	227	2	O4QSL5_9LECA	O4qsl5_ochrolechia
710	14	41.2	202	2	O5H3M5_XANOR	O5h3m5_xanthomonas	783	14	41.2	227	2	O4QSL4_9LECA	O4qsl4_pertusaria
711	14	41.2	203	2	O4QBG4_LEIMA	O4qbg4_leishmania	784	14	41.2	227	2	O4QSL3_9LECA	O4qsl3_pertusaria
712	14	41.2	203	2	Q4SEK7_TETNG	Q4sek7_tetradodon n	785	14	41.2	227	2	O4QSL1_9LECA	O4qsl1_pertusaria
713	14	41.2	204	1	DUS18_RAT	Q4exw7_tattus norv	786	14	41.2	227	2	O4QSK8_9LECA	O4qsk8_pertusaria
714	14	41.2	204	2	O5B220_EMBENI	O5b220_aspergillus	787	14	41.2	227	2	O4QSK6_9LECA	O4qsk6_pertusaria
715	14	41.2	204	2	O92472_NPVENI	O92472_bombyx mori	788	14	41.2	227	2	O4QSK4_9LECA	O4qsk4_pertusaria
716	14	41.2	205	1	Y115_NPVOF	O10354_oryzia pseu	789	14	41.2	227	2	O4QSK2_9LECA	O4qsk2_pertusaria
717	14	41.2	205	2	O94FP0_MIMCU	O94fpo_mimulus gut	790	14	41.2	227	2	O4QSV6_9LECA	O4qsv6_pertusaria
718	14	41.2	206	2	Q7Q606_ANOGA	Q7q606_anopheles g	791	14	41.2	227	2	O4QSV4_9LECA	O4qsv4_pertusaria
719	14	41.2	206	2	O8IHC4_DROME	O8ihc4_drosophila	792	14	41.2	227	2	O4QSV2_9LECA	O4qsv2_pertusaria
720	14	41.2	206	2	O8I934_DROME	O8i934_drosophila	793	14	41.2	227	2	O4QSV1_9LECA	O4qsv1_pertusaria
721	14	41.2	206	2	O6ZHP1_ORYSA	O6zhf1_oryza sativ	794	14	41.2	227	2	O4QSV5_9LECA	O4qsv5_pertusaria
722	14	41.2	206	2	O4SKY8_TETNG	O4skv8_tetradodon n	795	14	41.2	227	2	O4QSV4_9LECA	O4qsv4_pertusaria
723	14	41.2	206	2	O89183_BRETR	O89183_feline immu	796	14	41.2	227	2	O4QSV3_9LECA	O4qsv3_pertusaria
724	14	41.2	207	2	O4QSK0_9LECA	O4qsk0_pertusaria	797	14	41.2	227	2	O4QSV2_9LECA	O4qsv2_pertusaria
725	14	41.2	209	2	O94828_TETTH	O94828_tetrahymena	798	14	41.2	227	2	O4QSV1_9LECA	O4qsv1_pertusaria
726	14	41.2	209	2	QANDI1_9MICC	Qandil_athrobacte	799	14	41.2	227	2	O4QSH9_9LECA	O4qsh9_pertusaria
727	14	41.2	210	1	PRDX5_MOUSE	P99029_m_perotixed	800	14	41.2	227	2	O4QSH7_9LECA	O4qsh7_pertusaria
728	14	41.2	210	2	O4QSH4_9LECA	O4qsh4_pertusaria	801	14	41.2	227	2	O4QSH3_9LECA	O4qsh3_pertusaria
729	14	41.2	210	2	O4X8Z3_PLACH	O4x8z3_plasmodium	802	14	41.2	227	2	O4QSH2_9LECA	O4qsh2_pertusaria
730	14	41.2	210	2	Q7TLP3_NPVCF	Q7tlp3_choristoneu	803	14	41.2	227	2	O4QSH0_9LECA	O4qsh0_pertusaria
731	14	41.2	211	2	O4QSL7_9LECA	O4qsl7_ochrolechia	804	14	41.2	227	2	O4QSG7_9LECA	O4qsg7_pertusaria
732	14	41.2	211	2	O9D6X2_MOUSE	O9d6x2_mus musculu	805	14	41.2	227	2	O4QSG5_9LECA	O4qsg5_pertusaria
733	14	41.2	212	2	O4QSH5_9ASCO	O4qsh5_coccotrema	806	14	41.2	227	2	O4QSG4_9LECA	O4qsg4_pertusaria
734	14	41.2	212	2	O4QSHS_9LECA	O4qshs_pertusaria	807	14	41.2	227	2	O4QSG2_9LECA	O4qsg2_pertusaria
735	14	41.2	212	2	O5LA82_BACFN	O5la82_bacteroides	808	14	41.2	227	2	O4QSF8_9LECA	O4qsf8_pertusaria
736	14	41.2	212	2	O64OK1_BACFR	O64ok1_bacteroides	809	14	41.2	227	2	O4QSF4_9LECA	O4qsf4_pertusaria
737	14	41.2	213	2	O4QSK5_9LECA	O4qsk5_pertusaria	810	14	41.2	227	2	O4QSF3_9LECA	O4qsf3_pertusaria
738	14	41.2	213	2	O4QSH1_9LECA	O4qsh1_pertusaria	811	14	41.2	227	2	O4QSF1_9LECA	O4qsf1_pertusaria
739	14	41.2	213	2	O4QSF0_9LECA	O4qsf0_pertusaria	812	14	41.2	227	2	O4QSF9_9LECA	O4qsf9_varicellari
740	14	41.2	213	2	O4PSK8_USTMA	O4psk8_ustliago ma	813	14	41.2	227	2	O4QSF5_9LECA	O4qsf5_varicellari
741	14	41.2	213	2	O54E19_DICDI	O54e19_dicyosteli	814	14	41.2	227	2	O4QSF5_9LECA	O4qsf5_pertusaria
742	14	41.2	214	2	O5L6N1_CHLAB	O5l6n1_chlamydomphi	815	14	41.2	227	2	O4QSL2_9LECA	O4qsl2_pertusaria
743	14	41.2	214	2	O9F6W0_RHIFT	O9f6w0_myxobacteri	816	14	41.2	227	2	O4QSK1_9LECA	O4qsk1_pertusaria
744	14	41.2	214	2	O53846_MYCTU	O53846_myxobacteri	817	14	41.2	227	2	O9U4U2_DROME	O9u4u2_drosophila
745	14	41.2	214	2	Q7UL58_MYCBO	Q7ul58_mycobacteri	818	14	41.2	228	1	Q7JMT8_CAEEL	Q7jmt8_caenorhabdi
746	14	41.2	215	2	Q4WPS9_ASPFU	O4wfs9_aspergillus	819	14	41.2	229	1	Y4MB_RHISN	P55568_rhizobium s
747	14	41.2	215	2	O4PDI6_USTMA	O4pdi6_ustliago ma	820	14	41.2	229	2	O7Z9T2_9FUNG	O7z9t2_fungal endo
748	14	41.2	216	2	O5ICC4_EMBENI	O5icc4_emeritcellia	821	14	41.2	229	2	O7Z9T8_9FUNG	O7z9t8_fungal endo
749	14	41.2	216	2	O5AOK9_EMBENI	O5aok9_aspergillus	822	14	41.2	229	2	O7Z9U9_9FUNG	O7z9u9_fungal endo
750	14	41.2	216	2	O6KBI9_TKAVE	O6kbi9_trametes ve	823	14	41.2	229	2	O7Z9K0_9FUNG	O7z9k0_fungal endo
751	14	41.2	216	2	O4QSF2_9LECA	O4qsf2_pertusaria	824	14	41.2	229	2	O7Z9K1_9FUNG	O7z9k1_fungal endo
752	14	41.2	216	2	O6ZMZ4_HUMAN	O6zma4_homo sapien	825	14	41.2	229	2	O7Z9K2_9FUNG	O7z9k2_fungal endo
753	14	41.2	216	2	O6NAG6_HOPPA	O6nag6_rhodospseudo	826	14	41.2	229	2	O7X867_ORYSA	O7x867_oryza sativ
754	14	41.2	217	2	O6V276_SBACT	O6v276_eymbdiont ba	827	14	41.2	230	2	O5IUT3_MAGGR	O5iut3_magnaporthe
755	14	41.2	217	2	O5P3R3_AZOSS	O5p3r3_azocarcus sp	828	14	41.2	230	2	O7Z970_9FUNG	O7z970_fungal endo
756	14	41.2	217	2	O63PD8_BURPS	O63pd8_burkholderi	829	14	41.2	231	2	O7Z971_9FUNG	O7z971_fungal endo
757	14	41.2	217	2	O6ZEL4_BURMA	O6zel4_burkholderi	830	14	41.2	231	2	O80SR7_9ZZZZ	O80sr7_uncidentifie
758	14	41.2	217	2	O8PKP6_XANAC	O8pkp6_xanthomonas	831	14	41.2	232	2	O274Z3_DROYI	O274z3_drosophila
759	14	41.2	217	2	O80SR6_9ZZZZ	O80sr6_uncidentifie	832	14	41.2	233	2	O4RP76_TETNG	O4rp76_tetradodon n
760	14	41.2	217	2	O38020_9YTRU	O38020_potato viru	833	14	41.2	234	2	O69HP3_CIOIN	O69hp3_ciona intes
761	14	41.2	220	2	O4QSH6_9LECA	O4qsh6_pertusaria	834	14	41.2	234	2	O9VRR4_DROME	O9vrr4_drosophila

835	14	41.2	234	2	Q98AP0_RHIL0	Q98AP0_rhizobium 1	908	14	41.2	270	2	Q46004_CABEL	Q46004_caeonorthabdi
836	14	41.2	235	2	Q56PB1_9METZ	Q56PB1_uncultured	909	14	41.2	270	2	Q6FFC7_AC1AD	Q6FFC7_acetneobact
837	14	41.2	235	2	Q56PB3_9METZ	Q56PB3_uncultured	910	14	41.2	274	2	Q6JBK1_ZEADI	Q6JBK1_zea diploue
838	14	41.2	236	2	Q56P99_9METZ	Q56P99_uncultured	911	14	41.2	274	2	Q6TNB6_METEX	Q6TNB6_methylobact
839	14	41.2	236	2	Q4SZK8_TETNG	Q4SZK8_tetradodon n	912	14	41.2	274	2	Q6PNE6_XYLTP	Q6PNE6_xyliella fas
840	14	41.2	239	2	Q9Y831_9ASCO	Q9Y831_phoma sp. c	913	14	41.2	274	2	Q6TBC3_XYLTP	Q6TBC3_xyliella fas
841	14	41.2	239	2	Q5C403_SCHJA	Q5C403_schistosoma	914	14	41.2	274	2	Q6BLP2_XANAC	Q6BLP2_xanthomonas
842	14	41.2	239	2	Q7R376_GIALA	Q7R376_giardia lam	915	14	41.2	275	2	Q6I109_WHEAT	Q6I109_criticum ae
843	14	41.2	239	2	Q5BMW2_XANCA	Q5BMW2_xanthomonas	916	14	41.2	275	2	Q7W617_BORPA	Q7W617_bordeiella
844	14	41.2	239	2	Q5BMX0_XANCA	Q5BMX0_xanthomonas	917	14	41.2	275	2	Q7WIF9_BORBP	Q7WIF9_bordeiella
845	14	41.2	240	1	M810_AEATH	P92519_arabidopsis	918	14	41.2	275	2	Q7VWM2_BORPE	Q7VWM2_bordeiella
846	14	41.2	240	2	Q51EN0_9ARCH	Q51EN0_uncultured	919	14	41.2	276	2	Q6GV4_HOMAN	Q6GV4_homo sapien
847	14	41.2	240	2	Q51EN3_9ARCH	Q51EN3_uncultured	920	14	41.2	276	2	Q6NBD0_GLOMR	Q6NBD0_glossina mo
848	14	41.2	240	2	Q9Y864_ASPPA	Q9Y864_aspergillus	921	14	41.2	276	2	Q6JBKO_ZEADI	Q6JBKO_zea diploue
849	14	41.2	240	2	Q9HB81_HUMAN	Q9HB81_homo sapien	922	14	41.2	276	2	Q4UTS4_XANCP	Q4UTS4_xanthomonas
850	14	41.2	240	2	Q8T0D5_DROME	Q8T0D5_drosophila	923	14	41.2	276	2	Q5PEI1_SALTP	Q5PEI1_salmonella
851	14	41.2	241	2	Q9Y832_PENPA	Q9Y832_penicillium	924	14	41.2	276	2	Q6Z449_SALTP	Q6Z449_salmonella
852	14	41.2	241	2	Q9Y863_ASPPA	Q9Y863_aspergillus	925	14	41.2	276	2	Q6P9W2_XANCP	Q6P9W2_xanthomonas
853	14	41.2	242	2	Q51EM9_9ARCH	Q51EM9_uncultured	926	14	41.2	276	2	Q4S370_TETNG	Q4S370_tetradodon n
854	14	41.2	244	2	Q9Y100_9NEOP	Q9Y100_ectodeamia	927	14	41.2	277	2	Q9VPM8_DROME	Q9VPM8_drosophila
855	14	41.2	244	2	Q4U052_XANCP	Q4U052_xanthomonas	928	14	41.2	278	2	PLCB_MOUSE	Q8MK37_mus musculu
856	14	41.2	244	2	Q4P6V7_9RAC	Q4P6V7_uncultured	929	14	41.2	278	2	Q8MOQ2_DROME	Q8MOQ2_drosophila
857	14	41.2	244	2	Q6PHR5_XANAC	Q6PHR5_xanthomonas	930	14	41.2	278	2	Q92TN3_RHIME	Q92TN3_rhizobium m
858	14	41.2	245	2	Q6RK05_GIBZE	Q6RK05_gibberella	931	14	41.2	280	2	Q8WKG1_LOLMU	Q8WKG1_lotium mult
859	14	41.2	248	2	Q8P4J7_XANCP	Q8P4J7_xanthomonas	932	14	41.2	281	2	Q8JXM4_VYIRU	Q8JXM4_helictobis z
860	14	41.2	249	2	Q5AUV21_EMENI	Q5AUV21_aspergillus	933	14	41.2	282	2	Q7XL74_ORYSA	Q7XL74_oryza sativ
861	14	41.2	249	2	Q6ON68_CABRR	Q6ON68_caeonorthabdi	934	14	41.2	283	2	Q7D973_WYCTU	Q7D973_mycobacteri
862	14	41.2	249	2	Q16420_CABEL	Q16420_caeonorthabdi	935	14	41.2	285	2	Q7R3K3_GIALA	Q7R3K3_giardia lam
863	14	41.2	250	1	BTG4_MOUSE	Q70552_mus musculu	936	14	41.2	286	1	CKA6_RAT	P28233_rattus nov
864	14	41.2	250	2	Q61134_CABBR	Q61134_caeonorthabdi	937	14	41.2	286	2	Q6BRN2_MOUSE	Q6BRN2_mus musculu
865	14	41.2	250	2	Q16418_CABEL	Q16418_caeonorthabdi	938	14	41.2	286	2	Q4RV58_TETNG	Q4RV58_tetradodon n
866	14	41.2	250	2	Q16419_CABEL	Q16419_caeonorthabdi	939	14	41.2	287	2	Q8MVJ7_GASCI	Q8MVJ7_boltenia vi
867	14	41.2	250	2	Q5SND6_ORYSA	Q5SND6_oryza sativ	940	14	41.2	288	2	Q41937_GIBZE	Q41937_gibberella
868	14	41.2	250	2	Q5FUR7_GLUOX	Q5FUR7_glycine	941	14	41.2	288	2	Q9YV81_CABEL	Q9YV81_caeonorthabdi
869	14	41.2	250	2	Q5HWV9_CAMUR	Q5HWV9_campylobact	942	14	41.2	288	2	Q6H604_ORYSA	Q6H604_oryza sativ
870	14	41.2	250	2	Q925T9_MOUSE	Q925T9_mus musculu	943	14	41.2	289	2	Q7QV40_GIALA	Q7QV40_giardia lam
871	14	41.2	250	2	Q6MX24_MOUSE	Q6MX24_mus musculu	944	14	41.2	289	2	Q9VU78_DROME	Q9VU78_drosophila
872	14	41.2	251	2	Q4P139_USTMA	Q4P139_uscilleago ma	945	14	41.2	289	2	Q4V9J3_BRARE	Q4V9J3_brachydanto
873	14	41.2	251	2	Q4IAC0_GIBZE	Q4IAC0_gibberella	946	14	41.2	290	2	Q752M4_ASHGO	Q752M4_ashbya gos
874	14	41.2	253	2	Q6ZGX8_HUMAN	Q6ZGX8_homo sapien	947	14	41.2	290	2	Q9VX17_DROSOPH	Q9VX17_drosophila
875	14	41.2	253	2	Q7UY70_RHOBA	Q7UY70_rhodospirell	948	14	41.2	290	2	Q6GHV3_BRARE	Q6GHV3_brachydanto
876	14	41.2	254	2	Q6YUY0_ORYSA	Q6YUY0_oryza sativ	949	14	41.2	291	2	Q7XZP8_ORYSA	Q7XZP8_oryza sativ
877	14	41.2	254	2	Q55700_PRRSV	Q55700_porcine rep	950	14	41.2	291	2	Q6H5L8_ORYSA	Q6H5L8_oryza sativ
878	14	41.2	254	2	Q513H3_PRRSV	Q513H3_porcine rep	951	14	41.2	291	2	Q6ES25_ORYSA	Q6ES25_oryza sativ
879	14	41.2	254	2	Q6A521_PRRSV	Q6A521_porcine rep	952	14	41.2	292	2	Q8RIY7_MOUSE	Q8RIY7_mus musculu
880	14	41.2	254	2	Q84965_PRRSV	Q84965_porcine rep	953	14	41.2	293	2	Q7U3U8_SYNFX	Q7U3U8_synecococc
881	14	41.2	255	2	Q8N0W7_HUMAN	Q8N0W7_homo sapien	954	14	41.2	293	2	Q57124_PRRSV	Q57124_coho salmon
882	14	41.2	255	2	Q56P90_PRRSV	Q56P90_fungal sp.	955	14	41.2	293	2	Q6IWL1_XENLA	Q6IWL1_xenopus iae
883	14	41.2	255	2	Q4Q1V5_LEIMA	Q4Q1V5_leishmania	956	14	41.2	296	2	Q4IIN2_GIBZE	Q4IIN2_gibberella
884	14	41.2	257	1	CJ1095_HUMAN	Q55967_homo sapien	957	14	41.2	297	2	Q5ZN29_VYIRU	Q5ZN29_cotesia con
885	14	41.2	257	2	Q5Z9G6_ORYSA	Q5Z9G6_oryza sativ	958	14	41.2	298	2	Q99159_TRYFO	Q99159_trypanosoma
886	14	41.2	258	2	Q6ON66_CABBR	Q6ON66_caeonorthabdi	959	14	41.2	298	2	Q5VS69_ORYSA	Q5VS69_oryza sativ
887	14	41.2	258	2	Q9Y0E9_DROME	Q9Y0E9_drosophila	960	14	41.2	298	2	Q7U7H9_STNPK	Q7U7H9_synecococc
888	14	41.2	258	2	Q00946_9HYMN	Q00946_tetrahymena	961	14	41.2	299	2	Q5R0R8_IDILO	Q5R0R8_idiomarina
889	14	41.2	258	2	Q5PAGO_ANAMM	Q5PAGO_anaplasm	962	14	41.2	300	2	Q22513_POEYN	Q22513_glycine max
890	14	41.2	258	2	Q8GVG2_ADEBA	Q8GVG2_bovine aden	963	14	41.2	300	2	Q4ZTF5_PEEBY	Q4ZTF5_pseudomonas
891	14	41.2	259	2	Q7S668_NEUCR	Q7S668_neutrospora	964	14	41.2	300	2	Q9VGD9_ONCMY	Q9VGD9_oncorhynch
892	14	41.2	259	2	Q83212_TREPA	Q83212_treponema p	965	14	41.2	301	2	Q7V945_PROAM	Q7V945_prochloroco
893	14	41.2	261	2	Q94EL3_SORHL	Q94EL3_sorghum hal	966	14	41.2	301	2	Q6BN70_MOUSE	Q6BN70_mus musculu
894	14	41.2	261	2	Q5BK83_RAT	Q5BK83_rattus nov	967	14	41.2	301	2	Q6ELZ9_BRARE	Q6ELZ9_brachydanto
895	14	41.2	262	2	Q5DC20_SCHJA	Q5DC20_schistosoma	968	14	41.2	301	2	Q4VBS7_BRARE	Q4VBS7_brachydanto
896	14	41.2	264	2	Q27253_CABEL	Q27253_caeonorthabdi	969	14	41.2	302	2	Q7X546_9ACTO	Q7X546_actinoplan
897	14	41.2	264	2	Q08343_WYCTU	Q08343_mycobacteri	970	14	41.2	302	2	Q4SZH4_TETNG	Q4SZH4_tetradodon n
898	14	41.2	264	2	Q7VFO2_MYCBO	Q7VFO2_mycobacteri	971	14	41.2	304	2	Q4LET4_PILTI	Q4LET4_piltilaria
899	14	41.2	267	2	Q6E8T9_FUGRU	Q6E8T9_fugu rubrip	972	14	41.2	304	2	Q9CUT3_MOUSE	Q9CUT3_mus musculu
900	14	41.2	268	2	Q81110_WHEAT	Q81110_criticum ae	973	14	41.2	306	2	Q4UZB7_XANCP	Q4UZB7_xanthomonas
901	14	41.2	268	2	Q5H5W0_XANOR	Q5H5W0_xanthomonas	974	14	41.2	307	2	Q8PD32_XANCP	Q8PD32_xanthomonas
902	14	41.2	269	2	Q4Q5X9_LEIMA	Q4Q5X9_leishmania	975	14	41.2	307	2	Q94LD5_ORYSA	Q94LD5_oryza sativ
903	14	41.2	269	2	Q5GYN6_XANOR	Q5GYN6_xanthomonas	976	14	41.2	308	2	Q7R414_GIALA	Q7R414_giardia lam
904	14	41.2	269	2	Q6Z2Z2_BURPS	Q6Z2Z2_burkholderi	977	14	41.2	309	1	CF188_HUMAN	Q8H5C1_homo sapien
905	14	41.2	269	2	Q8KKG2_BURPS	Q8KKG2_burkholderi	978	14	41.2	309	1	CF188_MOUSE	Q8I100_mus musculu
906	14	41.2	269	2	Q8KKG3_BURPS	Q8KKG3_burkholderi	979	14	41.2	309	1	CF188_RAT	Q5W84_rattus nov
907	14	41.2	269	2	Q6Z2U6_BURPS	Q6Z2U6_burkholderi	980	14	41.2	309	2	Q9VZG9_DROME	Q9VZG9_drosophila



981	14	41.2	309	2	08P004_XMNAc	08P054_xanthomonas
982	14	41.2	311	2	08Y6XZ_9YIRU	08Y6XZ_iris yellow
983	14	41.2	311	2	091PB3_9YIRU	091PB3_iris yellow
984	14	41.2	312	2	016417_CAEEL	016417_caeornithabdi
985	14	41.2	315	2	Q7ONT6_GIALA	Q7rnt6_giardia lam
986	14	41.2	316	2	0680K2_ARATH	0680K2_arabidopsis
987	14	41.2	316	2	08L7U1_ARATH	08L7U1_arabidopsis
988	14	41.2	316	2	08L8A1_ARATH	08L8A1_arabidopsis
989	14	41.2	317	2	06FHE1_HOMAN	06Fhe1_homo sapiens
990	14	41.2	317	2	Q8WVZ0_TREIN	Q8wvz0_trichoplus
991	14	41.2	317	2	Q8BHY0_MOUSE	Q8bny0_mus musculus
992	14	41.2	318	2	069AN2_9SAUR	069an2_rhineci s
993	14	41.2	321	1	HEMH_WIGER	08d226_wiggleswort
994	14	41.2	321	2	Q7R109_GIALA	Q7r109_giardia lam
995	14	41.2	321	2	Q6NGE1_CORDI	Q6nge1_corynebacte
996	14	41.2	321	2	069AM6_9SAUR	069am6_xenochrophi
997	14	41.2	321	2	069AP5_9SAUR	069ap5_hemephros
998	14	41.2	321	2	069A97_9SAUR	069a97_platyceps r
999	14	41.2	321	2	069A02_9SAUR	069a02_platyceps r
1000	14	41.2	321	2	069A03_9SAUR	069a03_platyceps k

## ALIGNMENTS

RESULT 1	SGS3_DROME	STANDARD;	PRT;	307 AA.
ID	SGS3_DROME			
AC	P02840; Q9VTU2;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Salivary glue protein Sgs-3 precursor.			
GN	Name=SGS3; ORFNames=CG11720.			
OS	Drosophila melanogaster (fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
PN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEZLINE=83294545; PubMed=6411930;			
RA	Garfinkel M.D., Pruitt R.B., Meyerowitz E.M.;			
RT	"DNA sequences, gene regulation and modular protein evolution in the			
RT	Drosophila 68C glue gene cluster."			
RL	J. Mol. Biol. 168:765-789(1993).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=Berkeley;			
RX	MEZLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abbil J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,			
RA	Balcer J.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Bernan B.P., Bhandal D., Bolshakov S.,			
RA	Botkova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Dwyer A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Fowler K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischman W.,			
RA	Forster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaser K.,			
RA	Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Liou P., Lai Y., Levitsky A.C., Li J.H., Li Z., Liang Y., Lin X.,			
RA	Liao X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			

RA Metchukov G., Mikhina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusken D.R., Paclet J.M.,  
RA Palazolo M., Pletman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye Y., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RL "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
[3]  
RP GENOME REANNOTATION.  
RX MEDLINE=22426069; PubMed=12537572;  
RX Miya S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RX Rhee S.Y., Rhee S.Y., Rhee S.Y., Rhee S.Y., Rhee S.Y.,  
RA Hirdesky P., Huang Y., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter U., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RL "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RL systematic review.";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
[4]  
RP NUCLEOTIDE SEQUENCE OF 1-28.  
RX MEDLINE=88332966; PubMed=3138416;  
RX Martin C.H., Mayeda C.A., Meyerowitz E.M.;  
RA "Evolution and expression of the Sgs-3 glue gene of *Drosophila*.";  
RL J. Mol. Biol. 201:273-287(1988).  
[5]  
RP DEVELOPMENTAL STAGE.  
RX MEDLINE=94038699; PubMed=6223281;  
RA Huet F., Ruiz C., Richards G.;  
RT "Puffs and PCR: the in vivo dynamics of early gene expression during  
RT ecdysose responses in *Drosophila*.";  
RL Development 118:613-627(1993).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Salivary gland specific.  
CC -1- DEVELOPMENTAL STAGE: In the salivary glands of mid instar larvae  
CC levels dramatically increase during puff stage 1 at 98-106 hours  
CC of development. Levels remain constant and abundant in late larvae  
CC until puff stage 10, then decrease by stage 11.  
-----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
-----  
CC EMBL; X01918; CAA55994.1; -, Genomic DNA.  
DR EMBL; AE003544; AA050056.1; -, Genomic DNA.  
DR EMBL; X78392; CAA55154.1; -, Genomic DNA.  
DR PIR; A03329; GSFF3.  
DR Ensembl; CG11720; *Drosophila melanogaster*.  
KW FlyBase; FBgn0003373; Sgs3.  
KW Repeat; Signal.  
KW SIGNAL.  
FT CHAIN 1 23 Potential,  
FT 24 307 Salivary glue protein Sgs-3.  
SQ SEQUENCE 307 AA; 32196 MW; 45803DEBD16C418BC CRC64;  
  
Query Match 52.9%; Score 18; DB 1; Length 307;  
Best Local Similarity 16.7%; Pred. No. 4.9e-11;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
4 CXXXXXXXXXXC 15  
db 44 CTTTTTTTTTC 55



```

RESULT 2
08K340 MOUSE
ID 08K340_MOUSE PRELIMINARY; PRT; 347 AA.
AC 08K340;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE 463241106Rik protein (Fragment).
GN Name=463241106Rik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RC TISSUE=Mammary tumor metastasized to lung. MMTV-LTR/Mnt1 model.
RX Expression driven by an MMTV-LTR enhancer;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullanb S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallatton D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schneruch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RC TISSUE=Mammary tumor metastasized to lung. MMTV-LTR/Mnt1 model.
RX Expression driven by an MMTV-LTR enhancer;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028854; AAH28854.1; -; mRNA.
DR MGI; MGI:1925998; 463241106Rik.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR00159; RA.
DR PROSITE; PS50200; RA; 1.
FT NON TR
FT 1
SQ SEQUENCE 347 AA; 39161 MW; 306B40D38A14B3DE CRC64;
Query Match 52.9%; Score 18; DB 2; Length 347;
Best Local Similarity 16.7%; Pred. NO. 5.1e-11;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Cy 4 CXXXXXXXXXXC 15
Db 135 CSSTSSSTASSC 146

```

```

DE enriched library, clone:6030474M07 product:hypothetical Serine-rich
DE region/Aminoacyl-transfer RNA synthetases class-II containing protein,
DE full insert sequence.
GN Name=463241106Rik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi U., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bonb H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pezole G., Quackenbush J.,
RA Schirral L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barsh G.,
RA Blake J., Botfield D., Boujanna N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guenrich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seye T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welfz C., Whitlatch C., Wilting L.,
RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA The PANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoke S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitunui T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto K., Watsunoto H., Sakaguchi S., Ikegami T., Kasahagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.

```

RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RA Adachi J., Aizawa K., Hanigata T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Akimura T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Yamatsu M., Hayashizaki Y.,  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK031660; BAC27498.1; -, mRNA.  
 DR MGI; MGI:192598; 463241J06Rik.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR000159; RA.  
 DR SMART; SM00314; RA; 1.  
 DR PROSITE; PS50200; RA; 1.  
 KM Hypothetical protein.  
 SQ SEQUENCE 362 AA; 40934 MW; 32B1047ACC07B987 CRC64;

Query Match 52.9%; Score 18; DB 2; Length 362;  
 Best Local Similarity 16.7%; Pred. No. 5.2e-11;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15  
 Db 199 CSSTSSSTASSC 210

RESULT 4  
 ID 06AV38\_ORYSA PRELIMINARY; PRT; 367 AA.

AC 06AV38; 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DE 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DT Putative Cyclin.  
 GN Name=OSUNB0063J18.4;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Bell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,  
 RA Overton II L.L., Taitzin T., Kim M.M., Bera J.J., Jin S.S.,  
 RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsieh J., Blunt S.,  
 RA Vanaken S.S., Riedmiller S.B., Utterback T.T., Feldblyum T.V.,  
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,  
 RA White O., Salzberg S.L., Frazer C.M.; 3 BAC OSUNB0063J18 genomic sequence."  
 RT "Oryza sativa chromosome 3 BAC OSUNB0063J18 genomic sequence."  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.

RA Buell R.;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 CC -; SIMILARITY: Belongs to the cyclin family.  
 DR EMBL; AC107206; AAT77041.1; -, Genomic\_DNA.  
 DR Gramene; 06AV38; -.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0000074; P:regulation of cell cycle; IEA.  
 DR InterPro; IPR006670; Cyclin.  
 DR InterPro; IPR004367; Cyclin\_Cterm.  
 DR InterPro; IPR006671; Cyclin\_N.  
 DR Pfam; PF02984; Cyclin\_C; 1.  
 DR Pfam; PF00134; Cyclin\_N; 1.  
 DR SMART; SM00385; CYCLIN; 1.  
 KM Cyclin.  
 SQ SEQUENCE 367 AA; 38727 MW; 023BF8C0CDAD1D79 CRC64;

Query Match 52.9%; Score 18; DB 2; Length 362;  
 Best Local Similarity 16.7%; Pred. No. 5.2e-11;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 4 CXXXXXXXXXXC 15  
 Db 75 CSSTSSSTSSSSC 86

RESULT 5  
 ID 08BL43\_MOUSE PRELIMINARY; PRT; 584 AA.

AC 08BL43; 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length  
 DE enriched library, clone:B230384H21 product:hypothetical Serine-rich  
 DE region/Aminoacyl-transfer RNA synthetases class-II containing protein,  
 DE full insert sequence. (Fragment).  
 GN Name=463241J06Rik;  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning."  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1036/35055500;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehi Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi Y., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirnl L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,  
 RA Brownstein W.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontecki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 RN [4]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes."

RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600.  
 RA Shibata K., Itoh M., Atzawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Kono H., Akiyama J., Nishi K., Kitsuami T., Tashiro H., Itoh M.,  
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujisake K., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki K.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RA Adachi J., Atzawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Koda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK046433; BAC32726.1; - mRNA.  
 DR MGI: MGI:1925998; 463241J06R1K.  
 DR GO: GO:0007165; P: signal transduction; IEA.  
 DR InterPro: IPR001597; RA.  
 DR SMART: SM00314; RA; 1.  
 DR PROSITE: PS50200; RA; 1.  
 KM Hypothetical protein.  
 FT NON TR 1  
 SQ SEQUENCE 584 AA; 65345 MW; 941D03439114C6E CRC64;  
 Query Match 52.9%; Score 18; DB 2; Length 584;  
 Best Local Similarity 16.7%; Pred. No. 5.9e-11;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Qy 4 CXXXXXXXXXXC 15  
 Db 275 CSSTSSSTASGC 286  
 RESULT 6  
 OS BLP13\_PEA  
 ID QBLP13\_PEA PRELIMINARY; PRT; 730 AA.  
 AC QBLP13;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE ABA insensitive 3.  
 GN Name=ab13;  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
 NCBI\_TaxID=3888;  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Axillary bud;  
 RA Nakako A., Mori H.,  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB080195; BAC10553.1; - mRNA.  
 DR GO: GO:0003677; F: DNA binding; IEA.  
 DR GO: GO:0006355; P: regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR003340; TF\_B3.  
 DR Pfam: PF02362; B3; 1.

DR PROSITE: PS50863; B3; 1.  
 SQ SEQUENCE 730 AA; 80868 MW; 0A5B9A7350AB8B2 CRC64;  
 Query Match 52.9%; Score 18; DB 2; Length 730;  
 Best Local Similarity 16.7%; Pred. No. 6.2e-11;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Qy 4 CXXXXXXXXXXC 15  
 Db 55 CSSSSSSSSSSC 66  
 RESULT 7  
 OS ARATH  
 ID O9LQ12\_ARATH PRELIMINARY; PRT; 35 AA.  
 AC O9LQ12;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE F1504.2.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Ecker J.R.,  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Chouk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chiu J., Choi E.,  
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukhtarsky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaynsberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.,  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC007887; AAF79343.1; - genomic DNA.  
 SQ SEQUENCE 35 AA; 3663 MW; CB4370A9D6A5D1F4 CRC64;  
 Query Match 50.0%; Score 17; DB 2; Length 35;  
 Best Local Similarity 16.7%; Pred. No. 1.9e-09;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Qy 4 CXXXXXXXXXXC 15  
 Db 15 CSIASTAATSC 26  
 RESULT 8  
 OS GIALLA  
 ID Q24960\_GIALLA PRELIMINARY; PRT; 137 AA.  
 AC Q24960;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Tirophozolite surface protein 11 (Fragment).  
 GN Name=TSPI11/12;  
 OS Giardia lamblia (Giardia intestinalis).  
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
 NCBI\_TaxID=5741;  
 OX NCBI\_TaxID=5741;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC MEDLINE=98096860; PubMed=9435134;  
 RA Ey P.L., Mansouri M., Kulda J., Nohynkova E., Monis P.T.,  
 RA Andrews R.H., Mayrhofer G.,  
 RT "Genetic analysis of Giardia from hoofed farm animals reveals  
 RT atrodactyl-specific and potentially zoonotic genotypes.";  
 RL J. Eukaryot. Microbiol. 44:626-635(1997).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.

```

RA  EY P.L., Mansouri M., Kulda J., Nohynkova E., Monis P.T.,
RL  Andrews R.H., Mayrhofer G.;
DR  Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; U47631; AB05399.1; -, Genomic_DNA.
DR  InterPro; IPR006209; EGF_1like.
DR  InterPro; IPR005127; Giardia_VSP.
DR  Pfam; PF03302; VSP; 1.
DR  PROSITE; PS01186; EGF_2; 1.
FT  NON_TER
FT  NON_TER
SQ  SEQUENCE 137 AA; 13875 MW; 42346969B5ECF37D CRC64;

Query Match          50.0%; Score 17; DB 2; Length 137;
Best Local Similarity 16.7%; Pred. No. 2.7e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY  4 CXXXXXXXXXXC 15
DB  22 CASATARAATC 33

RESULT 9
Q24987_GIALA
ID  Q24987_GIALA PRELIMINARY; PRT; 137 AA.
AC  Q24987;
DT  01-NOV-1996 (TREMBLrel. 01, Created)
DT  01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT  01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE  Surface protein (Fragment).
GN  Name=VSP;
OS  Giardia lamblia (Giardia intestinalis).
OC  Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.
OX  NCBI_TaxID=5741;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  EY P.L., Darby J.M., Mayrhofer G.;
RL  Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
DR  EMBL; L16973; AB059198.1; -, Genomic_DNA.
DR  InterPro; IPR005127; Giardia_VSP.
DR  Pfam; PF03302; VSP; 1.
FT  NON_TER
FT  NON_TER
SQ  SEQUENCE 137 AA; 14148 MW; A340072153005D58 CRC64;

Query Match          50.0%; Score 17; DB 2; Length 137;
Best Local Similarity 16.7%; Pred. No. 2.7e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY  4 CXXXXXXXXXXC 15
DB  22 CTSTTARTATTC 33

RESULT 10
Q7XHS1_ORYSA
ID  Q7XHS1_ORYSA PRELIMINARY; PRT; 210 AA.
AC  Q7XHS1;
DT  01-OCT-2003 (TREMBLrel. 25, Created)
DT  01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT  01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE  2Fe-2S Iron-sulfur cluster protein-like.
GN  Name=P0477A12.11;
OS  Oryza sativa (Japonica cultivar-group).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC  Ehrhartoideae; Oryzaceae; Oryza.
OX  NCBI_TaxID=39947;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  Sasaki T., Matsumoto T., Katayose Y.;
RT  "Oryza sativa nupponbare (GA3) genomic DNA, chromosome 7, PAC
clone:P0477A12."
RL  Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

```

```

DR  EMBL; AP005190; BAC80058.1; -, Genomic_DNA.
DR  Gramene; Q7XHS1; -.
DR  GO; GO:0005489; F:electron transporter activity; IEA.
DR  GO; GO:0005506; F:iron ion binding; IEA.
DR  GO; GO:0046872; F:metal ion binding; IEA.
DR  GO; GO:0006118; P:electron transport; IEA.
DR  InterPro; IPR01041; Ferredoxin.
DR  InterPro; IPR012675; Ferredoxin_fold.
DR  Pfam; PF00111; Fer2; 1.
SQ  SEQUENCE 210 AA; 22629 MW; 6BCDC4F408B21E1E CRC64;

Query Match          50.0%; Score 17; DB 2; Length 210;
Best Local Similarity 16.7%; Pred. No. 3e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY  4 CXXXXXXXXXXC 15
DB  25 CAVATATTSSC 36

RESULT 11
Q9XY90_GIALA
ID  Q9XY90_GIALA PRELIMINARY; PRT; 234 AA.
AC  Q9XY90;
DT  01-NOV-1999 (TREMBLrel. 12, Created)
DT  01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT  01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE  Variant-specific surface protein (Fragment).
GN  Name=vsp417-4;
OS  Giardia lamblia (Giardia intestinalis).
OC  Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.
OX  NCBI_TaxID=5741;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=AQ-2;
RA  MEDLINE=99229863; PubMed=10215024; DOI=10.1016/S0166-6851(98)00183-2;
RA  EY P.L., Darby J.M., Mayrhofer G.;
RT  "A new locus (vsp417-4) belonging to the tsa417-like subfamily of
variant-specific surface protein genes in Giardia intestinalis."
RL  Mol. Biochem. Parasitol. 99:55-68(1999).
DR  EMBL; AF065600; A028789.1; -, Genomic_DNA.
DR  InterPro; IPR006212; Furin_repeat.
DR  InterPro; IPR005127; Giardia_VSP.
DR  InterPro; IPR006210; IEGF.
DR  Pfam; PF03302; VSP; 1.
DR  SMART; SM00181; EGF; 2.
DR  SMART; SM00261; FU; 1.
FT  NON_TER
FT  NON_TER
SQ  SEQUENCE 234 AA; 23565 MW; FDC75E280AF7D517 CRC64;

Query Match          50.0%; Score 17; DB 2; Length 234;
Best Local Similarity 16.7%; Pred. No. 3.1e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY  4 CXXXXXXXXXXC 15
DB  22 CTSATARAATC 33

RESULT 12
I18039_CABEL
ID  I18039_CABEL PRELIMINARY; PRT; 245 AA.
AC  I18039;
DT  01-JAN-1998 (TREMBLrel. 05, Created)
DT  01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT  01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE  Hypothetical protein T06C12.14.
GN  ORNames=T06C12.14;
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC  Rhabditidae; Pelodierinae; Caenorhabditis.
OX  NCBI_TaxID=6239;

```

```

RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
   investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81116; CAB03300.1; -; Genomic_DNA.
DR PIR; T24565; T24565.
DR Ensemble; T06C12.14; Caenorhabditis elegans.
DR WormBase; WBGene00011521; T06C12.14.
DR WormPep; T06C12.14; CE16364.
DR InterPro; IPR003582; SHKT.
DR Pfam; PF01549; SHKT; 3.
DR SMART; SM00254; SHKT; 4.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 245 AA; 25713 MW; BDE40649C966769A CRC64;

Query Match          50.0%; Score 17; DB 2; Length 245;
Best Local Similarity 16.7%; Pred. No. 3.1e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

CY 4 CXXXXXXXXXXC 15
DB 193 CASATTTTSSSTC 204

RESULT 13
Q5WQW8 CRYNE
ID Q5WQW8 CRYNE PRELIMINARY; PRT; 306 AA.
AC Q5WQW8;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Sec53p.
GN Name=SEC53;
OS Cryptococcus neoformans var. neoformans.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
ON NCBI_TaxID=40410;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RA Jambon G.;
RT "SEC53 encodes a putative phosphomannomutase.";
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY426175; AAR84595.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004615; F:phosphomannomutase activity; IEA.
DR GO; GO:0019307; P:mannose biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006379; HAD_SF_IIB.
DR InterPro; IPR005002; PMW_
DR Pfam; PF03332; PMW; 1.
DR TIGRfams; TIGR01484; HAD_SF_IIB; 1.
SQ SEQUENCE 306 AA; 33844 MW; FB9DB2CF1A3EDDEA CRC64;

Query Match          50.0%; Score 17; DB 2; Length 306;
Best Local Similarity 16.7%; Pred. No. 3.3e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

CY 4 CXXXXXXXXXXC 15
DB 5 CSSATTTTATTC 16

RESULT 14
Q6ASQ2 ORYSA
ID Q6ASQ2 ORYSA PRELIMINARY; PRT; 378 AA.
AC Q6ASQ2;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

```

```

DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein OSUNBA007732.21.
DN Name=OSUNBA007732.21;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
ON NCBI_TaxID=39947;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y., Hsing Y.-I. C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Zhao Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RT "Oryza sativa BAC OSUNBA007732 genomic sequence.";
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC148814; AAT93980.1; -; Genomic_DNA.
DR Gramene; Q6ASQ2; -.
KW Hypothetical protein.
SQ SEQUENCE 378 AA; 37703 MW; 342586481D07BF5B CRC64;

Query Match          50.0%; Score 17; DB 2; Length 378;
Best Local Similarity 16.7%; Pred. No. 3.5e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

CY 4 CXXXXXXXXXXC 15
DB 180 CVAAAAAATTAAC 191

RESULT 15
Q4LGT2_9BURK
ID Q4LGT2_9BURK PRELIMINARY; PRT; 396 AA.
AC Q4LGT2;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein.
ON ORFNames=Bcen2424DRAFT_0168;
OS Burkholderia cenocepacia H12424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cepacia complex.
ON NCBI_TaxID=331272;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RA US DOE Joint Genome Institute (JGI-PGF);
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Lartimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
   H12424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
   EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
   preliminary data.
DR EMBL; AALH01000139; EAM15270.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 396 AA; 41189 MW; C829063DC312B057 CRC64;

Query Match          50.0%; Score 17; DB 2; Length 396;
Best Local Similarity 16.7%; Pred. No. 3.5e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

QY 4 CXXXXXXXXXX 15  
 Db 18 CSSSDAAAAAAC 29

RESULT 16  
 ID Q8C8X1\_MOUSE PRELIMINARY; PRT; 429 AA.  
 Q8C8X1\_MOUSE  
 AC Q8C8X1;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
 DE Mus musculus adult retina cDNA, RIKEN full-length enriched library,  
 DE clone: A930006D11 product: hypothetical protein, full insert  
 DE sequence.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CS7BL/6J; TISSUE=Retina;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Mech. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CS7BL/6J; TISSUE=Retina;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batelov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barri G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohleutski S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CS7BL/6J; TISSUE=Retina;  
 RA The FANTOM Consortium,  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CS7BL/6J; TISSUE=Retina;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CS7BL/6J; TISSUE=Retina;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CS7BL/6J; TISSUE=Retina;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,  
 RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata N., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Yamamoto M., Hayashizaki Y.,  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK044299; BAC1860.1; -; mRNA.  
 DR Ensemble; ENSMUSG00000014198; Mus musculus.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR InterPro; IPR007087; ZnF\_C2H2.  
 DR InterPro; IPR003604; ZnF\_U1.  
 DR SMART; SM00355; ZnF\_C2H2; 3.  
 DR SMART; SM00451; ZnF\_U1; 3.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_3.  
 KM Hypothetical protein.  
 SQ SEQUENCE 429 AA; 44564 MW; 211C9F431F87FEF CRC64;

Query Match 50.0%; Score 17; DB 2; Length 429;  
 Best Local Similarity 16.7%; Pred. No. 3.6e-09;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXX 15  
 Db 181 CDAAASSSSSSC 192

RESULT 17  
 ID Q93H33\_STRAW PRELIMINARY; PRT; 540 AA.  
 AC Q93H33;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 13-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)  
 DE Transport system protein (putative oligopeptide ABC transporter  
 DE substrate-binding protein).  
 GN Name=opa2; OrderedLocustNames=SAV3149;  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.  
 NCBI\_Taxid=33903;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinoue M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NREL 8165;  
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbr820;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinohe M., Kikuchi H., Shiba T.,  
 RA Sasaki Y., Hattori M., Omura S.;  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism *Streptomyces avermitilis*,"  
 RL Nat. Biotechnol. 21:526-531 (2003).  
 DR EMBL: AB070952; BAB69358.1; -; Genomic DNA.  
 DR GO: 0005215; P: transporter activity; IEA.  
 DR GO: 0006810; P: transport; IEA.  
 DR InterPro: IPR000914; SBP\_bac\_5.  
 DR Pfam: PF00496; SBP\_bac\_5; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 540 AA; 58574 MW; 0052FEAD151B442C CRC64;

Query Match 50.0%; Score 17; DB 2; Length 540;  
 Best Local Similarity 16.7%; Pred. No. 3.8e-09;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15  
 Db 12 CAAAALAAATAC 23

RESULT 18  
 ID 09XTK3\_GIALA PRELIMINARY; PRT; 667 AA.  
 AC 09XTK3;  
 DT 01-NOV-1999 (TREMBLrel. 12; Created)  
 DT 01-NOV-1999 (TREMBLrel. 12; Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26; Last annotation update)  
 DE VS0417-3/A-II.  
 OS Name=vsp417-3/A-II;  
 GN Giardia lamblia (Giardia intestinalis).  
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.  
 CX NCBI\_TaxID=5741;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=BRIS/83/HEPU/136;  
 RA MEDLINE=99053029; PubMed=8836309;  
 RA By P.L., Darby J.M., Mayrhofer G.;  
 RT "Comparison of tsaa17-like variant surface protein (VSP)  
 RT genes in Giardia intestinalis and identification of a novel locus in  
 RT genetic group II isolates,"  
 RL Parasitology 117:445-455 (1998).  
 DR EMBL: AF033584; AAD03497.1; -; Genomic DNA.  
 DR HSSP: O16119; IEZG.  
 DR InterPro: IPR006212; Furin\_repeat.  
 DR InterPro: IPR005127; Giardia\_VSP.  
 DR InterPro: IPR006210; IEGF.  
 DR Pfam: PF03302; VSP; 1.  
 DR SMART: SMO0181; EGF; 2.  
 DR SMART: SMO0261; FU; 4.  
 SQ SEQUENCE 667 AA; 69123 MW; 77C64CF59441C0C CRC64;

Query Match 50.0%; Score 17; DB 2; Length 667;  
 Best Local Similarity 16.7%; Pred. No. 4e-09;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15  
 Db 448 CTSTTAAATAC 459

RESULT 19  
 ID 05B3H3\_EMENTI PRELIMINARY; PRT; 1031 AA.  
 AC 05B3H3;  
 DT 10-MAY-2005 (TREMBLrel. 30; Created)  
 DT 10-MAY-2005 (TREMBLrel. 30; Last sequence update)  
 DT 10-MAY-2005 (TREMBLrel. 30; Last annotation update)  
 DE Hypothetical protein.

GN ORFNames=AM4907.2;  
 OS Aspergillus nidulans FGSC A4.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiiales; Trichocomaceae; EmERICELLA.  
 CX NCBI\_TaxID=227321;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=FGSC A4;  
 RA Birren B., Nisbani C., Abouelleil A., Allen N., Anderson S.,  
 RA Archen H.M., Barna N., Bastien V., Bloom T., Boguslavsky L.,  
 RA Boukhvalter B., Butler J., Calvo S.B., Camarata J., Chang J.,  
 RA Chappel Y., Collamore A., Cook A., Cooke P., Corum B., Dearellano K.,  
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,  
 RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,  
 RA Gardina S., Gierre S., Graham L., Grand-Pierre N., Hafez N.,  
 RA Hagopian D., Hagos B., Hall J., Horon L., Hulme W., Iliev I.,  
 RA Jaffe D., Johnson R., Jones C., Kamel M., Kamet A., Karatas A.,  
 RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,  
 RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,  
 RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Menus L.,  
 RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,  
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,  
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,  
 RA Rachupka A., Ramasamy U., Raymond C., Rette R., Rise C., Rogov P.,  
 RA Roman J., Schauer S., Schupbach R., Seaman S., Severy P., Smirnov S.,  
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,  
 RA Talamas J., Testaye S., Theodore J., Topham K., Travers M.,  
 RA Vasiliiev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,  
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,  
 RA Zander E.;  
 RT "Genome Sequence of *Aspergillus nidulans*,"  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL: AACD0100084; EAA60985.1; -; Genomic DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1031 AA; 113626 MW; 6616443769CF2EF CRC64;

Query Match 50.0%; Score 17; DB 2; Length 1031;  
 Best Local Similarity 16.7%; Pred. No. 4.5e-09;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15  
 Db 743 CSTSSSASASTC 754

RESULT 20  
 ID MUC5B\_HUMAN STANDARD; PRT; 5703 AA.  
 AC 09HC84; 000447; 000573; 014985; 015494; 095291; 095451; Q14881;  
 AC 099552; 09UE28;  
 DT 28-FEB-2003 (Ref. 41; Created)  
 DT 28-FEB-2003 (Ref. 41; Last sequence update)  
 DT 13-SEP-2005 (Ref. 48; Last annotation update)  
 DE Mucin-5B precursor (Mucin 5 subtype B, tracheobronchial) (High  
 DE molecular weight salivary mucin MGL) (bublingual gland mucin).  
 GN Name=MUC5B; Synonyms=MUC5;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;  
 OC Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE OF 1-1594.  
 RA Chen Y., Di Y.P., Wu R.;  
 RT "Molecular cloning of the amino-terminal and 5'-flanking region of the  
 RT human MUC5B mucin gene,"  
 RN Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE OF 1-1325.  
 RX MEDLINE=9909274; PubMed=9790959; DOI=10.1006/dbrc.1998.9469;







```

FT CARBOHYD 5604 5604 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 5618 5618 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 5662 5662 N-linked (GlcNAc...) (Potential).
FT DISULFID 5594 5594 By similarity.
FT DISULFID 5622 5660 By similarity.
FT DISULFID 5626 5676 By similarity.
FT DISULFID 5643 5678 By similarity.
FT DISULFID 5645 5682 By similarity.
FT VARIANT 5137 5137 T->S (in dbSNP:2672788).
FT CONFLICT 34 34 /FTID=VAR_014123.
FT CONFLICT 95 100 PEGCN->LPCLK (in Ref. 2).
FT CONFLICT 104 104 S->C (in Ref. 2).
FT CONFLICT 142 142 E->K (in Ref. 1).
FT CONFLICT 225 225 R->S (in Ref. 2).
FT CONFLICT 330 331 PL->T (in Ref. 2).
FT CONFLICT 337 337 E->N (in Ref. 2).
FT CONFLICT 356 356 E->K (in Ref. 2).
FT CONFLICT 362 362 G->R (in Ref. 2).
FT CONFLICT 369 369 Missing (in Ref. 2 and 3).
FT CONFLICT 374 374 D->N (in Ref. 2).
FT CONFLICT 393 394 RT->TR (in Ref. 2).
FT CONFLICT 468 469 RK->GR (in Ref. 2).
FT CONFLICT 512 512 L->P (in Ref. 2).
FT CONFLICT 585 587 GAA->AH (in Ref. 3).
FT CONFLICT 601 601 A->S (in Ref. 3).
FT CONFLICT 628 629 DP->RS (in Ref. 2).
FT CONFLICT 633 633 F->L (in Ref. 2).
FT CONFLICT 676 676 A->P (in Ref. 3).
FT CONFLICT 701 701 R->P (in Ref. 3).
FT CONFLICT 752 752 E->K (in Ref. 2).

```

Query Match 50.0%; Score 17; DB 1; Length 5703;  
 Best Local Similarity 16.7%; Pred. No. 7.1e-09;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15  
 Db 5062 CTASATAAARC 5073

RESULT 21  
 ID 06EUA6\_ORYSA PRELIMINARY; PRT; 93 AA.  
 AC 06EUA6;  
 DT 25-OCT-2004 (TREMUREL. 28, Created)  
 DT 25-OCT-2004 (TREMUREL. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMUREL. 28, Last annotation update)  
 DE Hypothetical protein OJ1116\_C12.20.  
 GN Name-OJ1116\_C12.20;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Burkariota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Euphorbiaceae; Oryzae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP004134; BAD27763.1; -; Genomic\_DNA.  
 DR Gramene; 06EUA6; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 93 AA; 9912 MW; 10814C0A65B0F123 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 93;  
 Best Local Similarity 16.7%; Pred. No. 1.6e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15  
 Db 14 CTASAGDATSC 25

RESULT 22  
 ID 0850L9\_ORYSA PRELIMINARY; PRT; 100 AA.  
 AC 0850L9;  
 DT 01-JUN-2003 (TREMUREL. 24, Created)  
 DT 01-JUN-2003 (TREMUREL. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMUREL. 25, Last annotation update)  
 DE Putative topoisomerase.  
 GN Name-OSUNBA0032821.11;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Burkariota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Euphorbiaceae; Oryzae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Eastman A.P., Smith S.C., Bertin N., Liang C., Najjar F.Z., Pratt L.H.,  
 RA Cordonnier-Pratt M.-M.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF377947; A032310.1; -; Genomic\_DNA.  
 DR Gramene; 0850L9; -;  
 DR GO; GO:0016853; F:isomerase activity; IBA.  
 KW Isomerase.  
 SQ SEQUENCE 100 AA; 11026 MW; FA5BF1C9B0965999 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 100;  
 Best Local Similarity 16.7%; Pred. No. 1.6e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15  
 Db 60 CSSASAAAAPC 71

RESULT 23  
 ID 062UJ1\_BURMA PRELIMINARY; PRT; 100 AA.  
 AC 062UJ1;  
 DT 25-OCT-2004 (TREMUREL. 28, Created)  
 DT 25-OCT-2004 (TREMUREL. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMUREL. 28, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=BMA1578;  
 OS Burkholderia mallei (Pseudomonas mallei).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Burkholderiaceae; Burkholderia.  
 OX NCBI\_TaxID=13373;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA STRAIN=ATCC 23344;  
 RC PubMed=15377793; DOI=10.1073/pnas.040306101;  
 RA Nieman W.C., Deshaizer D., Kim H.S., Tettelin H., Nelson K.E.,  
 RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,  
 RA Daugherty S.C., Davidson T.D., DeBoy R.T., Dmitrov G., Dodson R.J.,  
 RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,  
 RA Madupu R., Mohammad Y., Nelson W.C., Radune C.M.,  
 RA Sarría S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,  
 RA Zafar N., Zhou L., Fraser C.M.;  
 RL "Structural flexibility in the Burkholderia mallei genome."  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).  
 DR EMBL; CP000010; AAU47930.1; -; Genomic\_DNA.  
 DR TIGR; BMA1578; -;  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 100 AA; 11000 MW; 89664B4C4C42D9F5 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 100;  
 Best Local Similarity 16.7%; Pred. No. 1.6e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15  
 Db 20 CLAAIAATATATC 31

```

RESULT 24
CBP1_AJECA
ID CBP1_AJECA STANDARD; PRT: 110 AA.
AC 042720;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Calcium-binding protein precursor (CBP).
GN Name=CBP1;
OS Ajellomyces capsulata (Histoplasma capsulatum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Oxygenales; Ajellomycesaceae; Ajellomyces.
OX NCBI_Taxid=5037;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RC STRAIN=G186AR;
RX MEDLINE=98149309; PubMed=9489665;
RA Batanghari J.W., Deepe G.S., Jr., Di Cera E., Goldman W.E.;
RT "Histoplasma acquisition of calcium and expression of CBP1 during
RT intracellular parasitism.";
RL Mol. Microbiol. 27:531-539(1998).
RN [2]
RP CHARACTERIZATION, AND MASS SPECTROMETRY.
RC STRAIN=G186AR;
RX MEDLINE=98053973; PubMed=9393824;
RA Batanghari J.W., Goldman W.E.;
RT "Calcium dependence and binding in cultures of Histoplasma
RT capsulatum.";
RL Infect. Immun. 65:5257-5261(1997).
CC -!- FUNCTION: Involved in calcium binding and uptake in yeast phase.
CC May be important for yeast survival during infection.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- DEVELOPMENTAL STAGE: Expressed in the yeast phase during
CC infection, and not in mycelia.
CC -!- PTM: May be glycosylated.
CC -!- MASS SPECTROMETRY: MW=7858.0; METHOD=Electrospray; RANGE=33-110;
CC NOTE=Ref. 2.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL: AF006209; AAC39354.1; -; Genomic DNA.
KW Calcium; Direct protein sequencing; Glycoprotein; Signal.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 110 Calcium-binding protein.
FT DISULFID 60 93 Potential.
FT DISULFID 65 76 Potential.
FT DISULFID 105 108 Potential.
SQ SEQUENCE 110 AA; 11014 MW; BD98F2FA1B4C77CE CRC64;

Query Match 47.1%; Score 16; DB 1; Length 110;
Best Local Similarity 16.7%; Pred. No. 1.7e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 65 CLSLSASSAAC 76

```

```

OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_Taxid=6182;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Han Z.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY810427; AAX26316.1; -; mRNA.
DR InterPro: IPR009081; ACP-like.
DR InterPro: IPR011989; ARM-like.
DR InterPro: IPR011992; EF-Hand_type.
KW Hypothetical protein.
SQ SEQUENCE 144 AA; 16762 MW; FF652713F976647B CRC64;

Query Match 47.1%; Score 16; DB 2; Length 144;
Best Local Similarity 16.7%; Pred. No. 1.8e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 15 CSSSYAVSSTSC 26

RESULT 26
Q9YBV0_AERPE
ID Q9YBV0_AERPE PRELIMINARY; PRT: 152 AA.
AC Q9YBV0;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein APE1499.
GN OrderedLocustNames=APE1499;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcates;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_Taxid=56636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KI;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Uin no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Horiyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix KI.";
RL DNA Res. 6:83-101(1999).
DR EMBL: BA000002; BAA80498.1; -; Genomic_DNA.
DR PIR: D72630; D72630.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 152 AA; 16241 MW; 4386978A0E6151EF CRC64;

Query Match 47.1%; Score 16; DB 2; Length 152;
Best Local Similarity 16.7%; Pred. No. 1.8e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15
DB 137 CSSTATTTLSTRC 148

RESULT 27
Q75LM9_ORYSA
ID Q75LM9_ORYSA PRELIMINARY; PRT: 161 AA.
AC Q75LM9;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative DNA topoisomerase (with alternative splicing).
GN Name=OSUNBA0047524.8;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC  Euphorbiaceae; Oryzaeae; Oryza.
OX  NCBI_TaxID=39947;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  Buehl C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA  Overton I.I.L., Tsetlin T., Kim M.M., Bera J.J., Jin S.S.,
RA  Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA  Vanaken S.S., Riedmiller S.B., Uferback T.T., Feldblum T.V.,
RA  Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA  White O., Salzberg S.L., Frazer C.M.;
RL  Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  NUCLEOTIDE SEQUENCE.
RA  Buehl R.;
RL  Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AC092556; AAR87258.1; -; Genomic_DNA.
DR  Gramene; Q75LM9; -.
DR  GO; GO:0005524; F:ATP binding; IEA.
DR  GO; GO:0003677; F:DNA binding; IEA.
DR  GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR  GO; GO:0016853; F:isomerase activity; IEA.
DR  GO; GO:0006285; P:DNA topological change; IEA.
DR  InterPro; IPR01991; wing_hlx_dna_bd.
DR  Pfam; PF04406; TP6A_N; 1.
DR  PRINTS; PR01550; TO6AFAMILY.
KM  DNA-binding; Isomerase; Transcription.
SQ  SEQUENCE 161 AA; 18051 MW; 41B1D72B71B78C3 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 161;
Best Local Similarity 16.7%; Pred. No. 1.9e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY  4 CXXXXXXXXXXC 15
DB  60 CSSSSSSAGSEC 71

RESULT 28
Q7PY44 ANOGA PRELIMINARY; PRT; 172 AA.
ID Q7PY44 ANOGA PRELIMINARY; PRT; 172 AA.
AC Q7PY44;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ENSANGP0000012250 (Fragment).
GN ORFNames=ENSANG0000009761;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygotha;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=PEST;
RC Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008987; EAA01398.1; -; Genomic_DNA.
DR InterPro; IPR001092; HLH_Basic.
DR Pfam; PF00010; HLH; 1.
DR PROSITE; PS50888; HLH; 1.
FT NON TER 1
SQ SEQUENCE 172 AA; 18464 MW; B857469A135A1864 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 172;
Best Local Similarity 16.7%; Pred. No. 1.9e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

QY  4 CXXXXXXXXXXC 15
DB  77 CSSSSSSAGSEC 88

RESULT 29
Q6ZD90 ORYSA PRELIMINARY; PRT; 176 AA.
ID Q6ZD90 ORYSA PRELIMINARY; PRT; 176 AA.
AC Q6ZD90;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein P0438H08.27.
GN Name=P0438H08.27;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AP004460; BAC99511.1; -; Genomic_DNA.
DR Gramene; Q6ZD90; -.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR01841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KM Hypothetical protein; Metal-binding; zinc; Zinc finger.
SQ SEQUENCE 176 AA; 18608 MW; 79FE9D42C740112 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 176;
Best Local Similarity 16.7%; Pred. No. 1.9e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY  4 CXXXXXXXXXXC 15
DB  96 CRRAAAAAASC 107

RESULT 30
Q55014 DICDI PRELIMINARY; PRT; 187 AA.
ID Q55014 DICDI PRELIMINARY; PRT; 187 AA.
AC Q55014;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DD80217903;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=AX4;
RC Eichinger L., Pachbat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugang R., Berriman M., Song J., Olsen R., Szafranski K., Xu O.,
RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankler A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pitcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,
RA Farbrother P., Desany B., Just B., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,

```

RA Loubegeed H., Mungall K., Oliver K., Price C., Quail M.A.,  
 RA Urushihara H., Hernandez J., Rabbindowitch E., Steffen D., Sanders M.,  
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,  
 RA Sngano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,  
 RA Shaulsky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,  
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
 RA Williams J., Dear P.H., Noegel A.A., Barrel B., Kuspa A.,  
 RT "The genome of the social amoeba *Dictyostelium discoideum*,"  
 RL Nature 0:0-0(2005).  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAFI01000042; EAL69047.1; -; Genomic\_DNA.  
 KM Hypothetical protein.  
 SQ SEQUENCE 187 AA; 20124 MW; 8D78DE804DFBC104 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 187;  
 Best Local Similarity 16.7%; Pred. No. 1.9e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
 DB 65 CSTRSSSSSFRSC 76

RESULT 31  
 ID 082LKO\_STRAW PRELIMINARY; PRT; 208 AA.  
 AC 082LKO; 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=SAV2010;  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=33903;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.,  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism *Streptomyces avermitilis*,"  
 RL Nat. Biotechnol. 21:526-531(2003).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,  
 RT "Genome sequence of an industrial microorganism *Streptomyces*  
 RT *avermitilis*: deducing the ability of producing secondary  
 RT metabolites,"  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 DR EMBL; BA000030; BAC69721.1; -; Genomic\_DNA.  
 KM Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 208 AA; 20876 MW; 4CDCSE8C111AF502 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 208;  
 Best Local Similarity 16.7%; Pred. No. 2e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
 DB 54 CSPAASASAAAC 65

RESULT 32  
 Q67VV9\_ORYSA

ID Q67VV9 ORYSA PRELIMINARY; PRT; 210 AA.  
 AC Q67VV9;  
 DT 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)  
 DE Hypothetical protein OSUNB0023122.2 (Hypothetical protein  
 DE P0530H05.17).  
 GN Name=OSUNB0023122.2; Synonyms=P0530H09.17;  
 OS Oryza sativa (japonica cultivar-group),  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC  
 RT clone:OSUNB0023122,"  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC  
 RT clone:P0530H05,"  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP004733; BAD37710.1; -; Genomic\_DNA.  
 DR EMBL; AP003541; BAD37385.1; -; Genomic\_DNA.  
 DR Gramene; Q67VV9; -;  
 KM Hypothetical protein.  
 SQ SEQUENCE 210 AA; 21575 MW; C5EB3FF3B9F4457 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 210;  
 Best Local Similarity 16.7%; Pred. No. 2e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 4 CXXXXXXXXXXC 15  
 DB 57 CSTRSSSSSRRRC 68

RESULT 33  
 ID Q5K9J0\_CRYNE PRELIMINARY; PRT; 214 AA.  
 AC Q5K9J0;  
 DT 10-MAY-2005 (TREMBlrel. 30, Created)  
 DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)  
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=CNK01770;  
 OS Cryptococcus neoformans var. neoformans JEC21.  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.  
 OX NCBI\_TaxID=214684;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=JEC21;  
 RA Loftus B., Amedeo P., Roncaglia P., Vamathavan J., Uterback T.,  
 RA Van Aken S., Fraser C.;  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).  
 RC STRAIN=JEC21;  
 RX PubMed=15653466; DOI=10.1126/science.1103773;  
 RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,  
 RA Vamathavan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,  
 RA Boeder I.E., Brent M.R., Chin R., Doering T.L., Donlin M.J.,  
 RA D'Souza C.A., Fox D.S., Gruberg V., Fu J., Fukushima M., Haas B.J.,  
 RA Huang J.C., Jambon G., Jones S.J.M., Koo H.L., Krzyzanski M.I.,  
 RA Kwon-Chung K.J., Lengeler K.B., Maiti R., Maiti R.A., Salzman S.L.,  
 RA Mathewson C.A., Mitchell T.G., Perera M., Riggs F.R., Salzberg S.L.,  
 RA Schein J.E., Shwartsbeyn A., Shin H., Shumway M., Specht C.A.,  
 RA Sun B.B., Tenney A., Uterback T.R., Wicks B.L., Wortman J.R.,  
 RA Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,

```

RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
RT Cryptococcus neoformans.";
RL Science 307:1321-1324(2005).
DR EMBL; AE017351; AAM46337.1; -; Genomic_DNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005732; C:small nuclear ribonucleoprotein complex; IEA.
DR GO; GO:0006397; P:mRNA processing; IEA.
DR InterPro; IPR001163; LSM_snrnp_core.
DR Pfam; PF01423; LSM; 1.
DR SMART; SM00651; Sm; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 214 AA; 23828 MW; 9DEFCE484BCE42 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 214;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 9 CYASTSSVASC 20

RESULT 34
Q55K09 CRYNE PRELIMINARY; PRT; 215 AA.
ID 055K09 CRYNE PRELIMINARY; PRT; 215 AA.
AC 055K09;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=CNBK1770;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC Preliminary data.
CC EMBL; AAEY0100052; EAL18156.1; -; Genomic_DNA.
CC Hypothetical protein.
SQ SEQUENCE 215 AA; 24014 MW; 1CEDEFCE484BCE CRC64;

Query Match 47.1%; Score 16; DB 2; Length 215;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 9 CYASTSSVASC 20

RESULT 35
Q5K919 CRYNE PRELIMINARY; PRT; 215 AA.
ID 05K919 CRYNE PRELIMINARY; PRT; 215 AA.
AC 05K919;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNK01770;
OS Cryptococcus neoformans var. neoformans JEC21.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=214684;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B., Amedeo P., Roncaglia P., Vamathavan J., Uterback T.,
RA Van Aken S., Fraser C.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=JEC21;
RX PubMed=15653466; DOI=10.1126/science.1103773;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathavan J., Miranda M., Anderson I.W., Fraser J.A., Allen J.E.,
RA Bosdet I.B., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Jambon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
RA Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
RA Matheson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,
RA Schein J.E., Shwartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA Sub B.B., Tenney A., Uterback T.R., Wickes B.L., Wortman J.R.,
RA Wye N.H., Kronstad J.W., Lodge J.K., Hettman J., Davis R.W.,
RA Fraser C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
RT Cryptococcus neoformans.";
RL Science 307:1321-1324(2005).
DR EMBL; AE017351; AAM46338.1; -; Genomic_DNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005732; C:small nuclear ribonucleoprotein complex; IEA.
DR GO; GO:0006397; P:mRNA processing; IEA.
DR InterPro; IPR001163; LSM_snrnp_core.
DR Pfam; PF01423; LSM; 1.
DR SMART; SM00651; Sm; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 215 AA; 24014 MW; 1CEDEFCE484BCE CRC64;

Query Match 47.1%; Score 16; DB 2; Length 215;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 9 CYASTSSVASC 20

RESULT 36
Q4OSJ9 9LECA PRELIMINARY; PRT; 227 AA.
ID 04OSJ9 9LECA PRELIMINARY; PRT; 227 AA.
AC 04OSJ9;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Putative non-reducing polyketide synthase (Fragment).
OS Pertusaria coronata.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Lecanoromycetes;
OC Pertusariales; Pertusariaceae; Pertusaria.
OX NCBI_TaxID=232756;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15927215; DOI=10.1016/j.phytochem.2005.04.014;
RA Schmitt I., Martin M.P., Kautz S., Lumbach H.T.;
RT "Diversity of non-reducing polyketide synthase genes in the
RT Pertusariales (lichenized Ascomycota): A phylogenetic perspective.";
RL Phytochemistry 66:1241-1253(2005).
DR EMBL; AY918748; AAY00086.1; -; Genomic_DNA.
DR InterPro; IPR000794; ketoacyl_synth.
DR Pfam; PF00103; ketoacyl_synth; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR TRANSFERASE.
FT NON TER 1
FT NON TER 227
SQ SEQUENCE 227 AA; 24445 MW; 4D57BB346B626A8C CRC64;

Query Match 47.1%; Score 16; DB 2; Length 227;
Best Local Similarity 16.7%; Pred. No. 2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

QY 4 CXXXXXXXXXXC 15  
DB 95 CESSPAIATAC 106

RESULT 37  
Q7JNB5\_GIALA PRELIMINARY; PRT; 230 AA.  
ID Q7JNB5;  
AC Q7JNB5;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Variant-specific surface protein type 4 tsap1/tsa417-like  
DE (Fragment).  
GN Name=vsp417-4;  
OS Giardia lamblia (Giardia intestinalis).  
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.  
OX NCBI\_TaxID=5741;  
RN [1]

RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=99229863; PubMed=10215024; DOI=10.1016/S0166-6851(98)00183-2;  
RA Ey P.L., Darby J.M., Mayrhofer G.;  
RT "A new locus (vsp417-4) belonging to the tsa417-like subfamily of  
RT variant-specific surface protein genes in Giardia intestinalis.";  
RL Mol. Biochem. Parasitol. 99:55-68(1999).  
DR EMBL: U89153; AAD05041.1; -; Genomic\_DNA.  
DR InterPro: IPR006212; Furin\_repeat.  
DR InterPro: IPR005127; Giardia\_VSP.  
DR Pfam: PF03302; VSP. 1.  
DR SMART: SM00181; EGF. 2.  
DR SMART: SM00261; FU; 1.  
FT NON\_TER 1  
FT NON\_TER 230  
SQ SEQUENCE 230 AA; 23336 MW; 45E22ECD1C00B4F1 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 230;  
Best Local Similarity 16.7%; Pred. No. 2e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 22 CSSTTVRAAATC 33

RESULT 38  
Q8LD53\_ARATH PRELIMINARY; PRT; 234 AA.  
ID Q8LD53;  
AC Q8LD53;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE bHLH transcription factor, putative.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
OC rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]

RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22088475; PubMed=12093376;  
RA Haas B.U., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,  
RA Feldmann K.A., Flavell R.B., White O., Salberg S.L.;  
RT "Full-length messenger RNA sequences greatly improve genome  
RT annotation.";  
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).  
RN [2]

RP NUCLEOTIDE SEQUENCE.  
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
RA Feldmann K.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

DR EMBL: AY086197; AA064276.1; -; mRNA.  
DR HSSP: P36956; 1AM9.  
DR GO: GO:0005634; C:nucleus; IEA.  
DR GO: GO:0030528; F:transcription regulator activity; IEA.  
DR GO: GO:0045449; P:regulation of transcription; IEA.  
DR InterPro: IPR001092; HLH\_basic.  
DR Pfam: PF00010; HLH; 1.  
DR SMART: SM00353; HLH; 1.  
DR PROSITE: PSS0868; HLH; 1.  
SQ SEQUENCE 234 AA; 25470 MW; 9EBEA7D2E44F21D5 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 234;  
Best Local Similarity 16.7%; Pred. No. 2e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
DB 67 CESSSATSSXAC 78

RESULT 39  
Q9FH37\_ARATH PRELIMINARY; PRT; 234 AA.  
ID Q9FH37;  
AC Q9FH37;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
DE Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K5F14  
DE (Hypothetical protein At5g54680, K5F14.2) (Putative bHLH transcription  
DE factor) (Hypothetical protein At5g54680).  
GN Name=At5g54680/K5F14.2; Synonym=At5g54680;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
OC rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]

RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20181125; PubMed=10718197;  
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
RA Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC  
RT clones.";  
RL DNA Res. 7:31-63(2000).  
RN [2]

RP NUCLEOTIDE SEQUENCE.  
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,  
RA Palm C.U., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamaya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]

RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22598051; PubMed=12679534; DOI=10.1093/molbev/meg088;  
RA Heim M.A., Jakoby M., Werber M., Martin C., Weisshaar B., Bailey P.C.;  
RT "The basic helix-loop-helix transcription factor family in plants: a  
RT genome-wide study of protein structure and functional diversity.";  
RL Mol. Biol. Evol. 20:735-747(2003).  
RN [4]

RP NUCLEOTIDE SEQUENCE.  
RA Jakoby M.J., Heim M.A., Bailey P., Martin C., Weisshaar B.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
RN [5]

RP NUCLEOTIDE SEQUENCE.  
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,  
RA Palm C.U., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamaya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.;

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.  
 DR EMBL: AB022214; BAB09914.1; -; Genomic\_DNA.  
 DR EMBL: AY054585; AAK96776.1; -; mRNA.  
 DR EMBL: AF488629; AAM10964.1; -; mRNA.  
 DR EMBL: BT002189; AAN72200.1; -; mRNA.  
 DR HSPSP: P36956; IAM9.  
 DR GO: GO:0045449; P:regulation of transcription; TAS.  
 DR InterPro: IPR001092; HLH\_basic.  
 DR Pfam: PF00010; HLH; 1.  
 DR SMART: SM00353; HLH; 1.  
 DR PROSITE: PS50888; HLH; 1.  
 KM Hypothetical protein.  
 SQ SEQUENCE 234 AA; 25488 MW; 987EA11424F21D1 CRC64;  
 Query Match 47.1%; Score 16; DB 2; Length 234;  
 Best Local Similarity 16.7%; Pred. No. 2e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Oy 4 CXXXXXXXXXXC 15  
 Db 67 CESSATSSKXC 78

RESULT 40  
 ID 016421\_CABEL PRELIMINARY; PRT; 244 AA.  
 AC 016421;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein T05B4.12.  
 GN ORFNames=T05B4.12;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RG The C. elegans sequencing consortium;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology";  
 RL Science 282:2012-2018(1998).  
 DR EMBL: AF016445; AAC69052.1; -; Genomic\_DNA.  
 DR PIR: T31838; T31838.  
 DR Ensembl: T05B4.12; Caenorhabditis elegans.  
 DR WormBase: WBGene00020243; T05B4.12.  
 DR WormPep: T05B4.12; CE13215.  
 DR InterPro: IPR003582; SHKT.  
 DR Pfam: PF01549; SHTK; 3.  
 DR SMART: SM00254; SHKT; 4.  
 KM Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 244 AA; 25347 MW; 36362492AC86B1C9 CRC64;  
 Query Match 47.1%; Score 16; DB 2; Length 244;  
 Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Oy 4 CXXXXXXXXXXC 15  
 Db 203 CTSTADSSSTSC 214

RESULT 41  
 ID 0966B6\_CABEL PRELIMINARY; PRT; 247 AA.  
 AC 0966B6;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein.

GN ORFNames=Y46H3D.8;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RG The C. elegans sequencing consortium;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology";  
 RL Science 282:2012-2018(1998).  
 DR EMBL: AC006777; AAK72311.1; -; Genomic\_DNA.  
 DR Ensembl: Y46H3D.8; Caenorhabditis elegans.  
 DR InterPro: IPR003582; SHKT.  
 DR Pfam: PF01549; SHTK; 3.  
 DR SMART: SM00254; SHKT; 4.  
 KM Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 247 AA; 25215 MW; FF3CD3BFD0DDE5A7 CRC64;  
 Query Match 47.1%; Score 16; DB 2; Length 247;  
 Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Oy 4 CXXXXXXXXXXC 15  
 Db 206 CTSTADSSSNC 217

RESULT 42  
 ID 016424\_CABEL PRELIMINARY; PRT; 248 AA.  
 AC 016424;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein T05B4.3.  
 GN ORFNames=T05B4.3;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RG The C. elegans sequencing consortium;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology";  
 RL Science 282:2012-2018(1998).  
 DR EMBL: AF016445; AAC69054.1; -; Genomic\_DNA.  
 DR PIR: T31841; T31841.  
 DR Ensembl: T05B4.3; Caenorhabditis elegans.  
 DR WormBase: WBGene00020237; T05B4.3.  
 DR WormPep: T05B4.3; CE13197.  
 DR InterPro: IPR003582; SHKT.  
 DR Pfam: PF01549; SHTK; 3.  
 DR SMART: SM00254; SHKT; 4.  
 KM Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 248 AA; 25659 MW; 79B281AFB7ED3AFC CRC64;  
 Query Match 47.1%; Score 16; DB 2; Length 248;  
 Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Oy 4 CXXXXXXXXXXC 15  
 Db 207 CTSTADSSSNC 218

RESULT 43  
 ID 09XV90\_CABEL

ID O9XV90 CAEEL PRELIMINARY; PRT; 257 AA.  
 AC O9XV90;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Hypothetical protein F16H6.3.  
 OS Name=F16H6.3; ORNames=F16H6.3;  
 GN Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderiinae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 RN [1]  
 RP NCLEOTIDE SEQUENCE.  
 RC STRAIN=Bristol N2;  
 RA Matthews L.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=9906613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;  
 RG The C. elegans sequencing consortium;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z81506; CAB04129.1; -; Genomic\_DNA.  
 DR PIR; T21029; T21029.  
 DR Ensemble; F16H6.3; Caenorhabditis elegans.  
 DR Normbase; WBSene0008893; F16H6.3.  
 DR WormPep; F16H6.3; C818599.  
 DR InterPro; IPR003582; SMT.  
 DR Pfam; PF01549; SMTK. 3.  
 DR SMART; SMO0254; SMTK. 4.  
 KM Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 257 AA; 27646 MW; CA204F1AC00EB87 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 257;  
 Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 QY 4 CXXXXXXXXXC 15  
 DB 195 CTNATTTTQASC 206  
 RESULT 44  
 O8CAZ5 MOUSE  
 ID O8CAZ5\_MOUSE PRELIMINARY; PRT; 262 AA.  
 AC O8CAZ5;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Mus musculus 6 days neonate skin cDNA, RIKEN full-length enriched  
 DE library, clone:A030004J08 product:hypothetical protein, full insert  
 DE sequence. (Fragment).  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP NCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Skin;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning."  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP NCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Skin;  
 RX MEDLINE=21086660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai U., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Hara A., Fukunishi Y., Komno H., Adachi U., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,

RA Saito T., Okazaki Y., Gojobori T., Bond H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Cavatoni T.,  
 RA Fleischmann W., Gaasterland M., Gissi C., King B., Kochava H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Peole G., Quackenbush J.,  
 RA Schiraldi L.M., Scabini F., Suzuki R., Tomita M., Wagner U., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barin G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guastich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,  
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP NCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Skin;  
 RA The PANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP NCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Skin;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes."  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP NCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Skin;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagasaka S., Sasaki N., Carninci P.,  
 RA Komno H., Akiyama U., Nishi K., Kiteunui T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Matsumi M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multiplexed capillary sequencer."  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP NCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Skin;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Harada A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirota K., Hirose T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Komno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takagawa-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK037168; BAC29730.1; -; mRNA.  
 DR Ensemble; ENSMUSG0000050239; Mus musculus.  
 KW Hypothetical protein.  
 FT NON TER  
 SQ SEQUENCE 262 AA; 29111 MW; 8D441316CD4432B8 CRC64;  
 Query Match 47.1%; Score 16; DB 2; Length 262;  
 Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;



```

Qy 4 CXXXXXXXXXXC 15
Db 139 CTTATRSASXAC 150

RESULT 45
Q7PRO7 ANOGA PRELIMINARY; PRT; 269 AA.
ID 07PRO7 ANOGA PRELIMINARY; PRT; 269 AA.
AC 07PRO7 ANOGA PRELIMINARY; PRT; 269 AA.
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ENSANGP0000001657 (Fragment).
GN ORFNames=ENSANGC0000001387;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (Apr-2004) to the EMBL/Genbank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA0100847; BAA06779.3; -; Genomic_DNA.
FT NON TER 1 269
FT NON TER 1 269
SQ SEQUENCE 269 AA; 24228 MW; 25BBFF71FD71F1F2 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 269;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 114 CASCSSTSSC 125

RESULT 46
Q6ZUB7 HUMAN PRELIMINARY; PRT; 278 AA.
ID 06ZUB7 HUMAN PRELIMINARY; PRT; 278 AA.
AC 06ZUB7 HUMAN PRELIMINARY; PRT; 278 AA.
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypothetical protein FLJ43840.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RG Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RG Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RG Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RG Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RG Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RG Murakawa K., Kanohori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RG Kawakami B., Suzuki Y., Sugano S., Nagahari K., Maeno Y., Nagai K.,
RG Iwagaki T.;
RL Submitted (JUL-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK125828; BAC6309.1; -, mRNA.

```

```

SQ SEQUENCE 278 AA; 2768 MW; B0C8757FC1DAE2C0 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 278;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 35 CTAAGSSSLSTC 46

RESULT 47
Q8PFE6 XANAC PRELIMINARY; PRT; 285 AA.
ID 08PFE6 XANAC PRELIMINARY; PRT; 285 AA.
AC 08PFE6 XANAC PRELIMINARY; PRT; 285 AA.
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Endonuclease.
GN OrderedLocusNames=XAC4037;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Faria L.R.,
RA Quagga R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F., Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camaroto G., Camaroto J., Chamergo F.,
RA Ciapina L.P., Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorriy H., Faria J.B., Ferreira A.J.B., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Seta J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamara R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AE012052; AAM3872.1; -, Genomic_DNA.
DR HSSP; P13717; 1G8T.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001604; Endonuclease.
DR Pfam; PF01223; Endonuclease_NS; 1.
DR SMART; SM00477; NUC; 1.
DR PROSITE; PS01070; NUCLEASE_NON_SPEC; UNKNOWN_1.
KW Complete proteome; Endonuclease.
SQ SEQUENCE 285 AA; 30358 MW; 5581A0E9A2057BB CRC64;

Query Match 47.1%; Score 16; DB 2; Length 285;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CXXXXXXXXXXC 15
Db 51 CTSVAATATTC 62

RESULT 48
Q6ZQ01 HUMAN PRELIMINARY; PRT; 286 AA.
ID 06ZQ01 HUMAN PRELIMINARY; PRT; 286 AA.
AC 06ZQ01 HUMAN PRELIMINARY; PRT; 286 AA.
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

```

DE Hypothetical protein FLJ46328.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;  
 OC Homo.  
 NCBI\_TaxID=9606;  
 RN  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Testis;  
 RA Nihomaya K., Magatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,  
 RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,  
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,  
 RA Otsuka T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
 RA Kawai-Hio Y., Sato K., Nishikawa T., Kimura K., Yamashita H.,  
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,  
 RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,  
 RA Suzuki Y., Sugano S., Nagahari K., Masuko Y., Nagai K., Isogai T.,  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK128833; BAC87633.1; -; mRNA.  
 SQ SEQUENCE 286 AA; 30415 MW; B6CEFA76737B8C CRC64;

Query Match 47.1%; Score 16; DB 2; Length 286;  
 Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
 DB 231 CTAAGSSSLSTC 242

RESULT 49  
 Q9F934\_BIFAD  
 ID Q9F934\_BIFAD PRELIMINARY; PRT; 287 AA.  
 AC Q9F934;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Sugar binding protein Sbp.  
 GN Name=sbp;  
 OS Bifidobacterium adolescentis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;  
 CC Bifidobacteriaceae; Bifidobacterium.  
 NCBI\_TaxID=1680;  
 RN  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=INT57;  
 RA Park M.S., Yoon H.J., Ji G.E.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF213175; AAG31696.1; -; Genomic DNA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro: IPR006059; SBP\_bac\_1.  
 DR Pfam: PF01547; SBP\_bac\_1;  
 SQ SEQUENCE 287 AA; 32054 MW; 11B9382AE39B86DB CRC64;

Query Match 47.1%; Score 16; DB 2; Length 287;  
 Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
 DB 48 CAAIMAASAAAC 59

RESULT 50  
 Q89DY5\_BRAJA  
 ID Q89DY5\_BRAJA PRELIMINARY; PRT; 297 AA.  
 AC Q89DY5;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE B117302 protein.  
 GN OrderedLocustNames=b117302;

OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 CC Bradyrhizobiaceae; Bradyrhizobium.  
 NCBI\_TaxID=375;  
 RN  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=USDA 110;  
 RX MEDLINE=2248498; PubMed=12597275;  
 RA Kaneko T., Nakamura Y., Sato S., Minamide K., Uehimi T.,  
 RA Sasaoka S., Watanabe A., Idegawa K., Iriuchi M., Kawashima K.,  
 RA Kohara M., Matsumoto M., Shimo S., Tsunoka H., Wada T., Yamada M.,  
 RA Tabata S.;  
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
 Bradyrhizobium japonicum USDA110.";  
 RL DNA Res. 9:189-197(2002).  
 DR EMBL; BA000040; BACS2567.1; -; Genomic DNA.  
 KW Complete proteome.  
 SQ SEQUENCE 297 AA; 32798 MW; E8E4D21866645601 CRC64;

Query Match 47.1%; Score 16; DB 2; Length 297;  
 Best Local Similarity 16.7%; Pred. No. 2.2e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 4 CXXXXXXXXXXC 15  
 DB 250 CSATHSSDAAC 261

Search completed: January 4, 2006, 16:09:50  
 Job time : 143.548 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 : Search time 54.7478 Seconds  
(without alignments)  
64,204 Million cell updates/sec

Title: US-09-932-322-9  
Perfect score: 24  
Sequence: 1 CXXXXXXC 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues  
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : A Geneseq\_21:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*
- 9: geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	75.0	9	8	ADQ13022 Hepatitis
2	18	75.0	9	8	ADQ13021 Hepatitis
3	18	75.0	11	2	AAW65943 Molecule
4	18	75.0	11	2	AAW65944 Molecule
5	18	75.0	15	9	ADV23688 HBV Immun
6	18	75.0	15	9	ADV23687 HBV Immun
7	18	75.0	24	4	ABR42273 Peptide #
8	18	75.0	24	4	AAW36081 Peptide #
9	18	75.0	24	4	AAW75973 Human bon
10	18	75.0	24	4	AAW63159 Human bra
11	18	75.0	24	4	ABG57702 Human liv
12	18	75.0	69	4	ABR68695 Drosophi1
13	18	75.0	69	8	ADP30715 Human sec
14	18	75.0	91	4	ABR39058 Peptide #
15	18	75.0	91	4	AAW32549 Peptide #
16	18	75.0	91	4	AAW72280 Human bon
17	18	75.0	91	4	ABG53976 Human bra
18	18	75.0	91	4	ABG42105 Human liv
19	18	75.0	93	5	ADP30859 Human pep
20	18	75.0	93	7	ABO83167 Pseudomon
21	18	75.0	99	8	ADP30709 Human sec
22	18	75.0	107	7	ABO74878 Pseudomon
23	18	75.0	115	7	ABO73394 Pseudomon
24	18	75.0			

25	18	75.0	122	7	ADP04777 Bacterial
26	18	75.0	135	8	ADP30703 Human sec
27	18	75.0	144	7	ADP05730 Pseudomon
28	18	75.0	144	8	ADP31453 Human sec
29	18	75.0	144	8	ADP90450 Plant ful
30	18	75.0	150	6	ABU45351 Protein e
31	18	75.0	153	5	ABR41043 Human ORF
32	18	75.0	153	5	ADP09251 Human ORF
33	18	75.0	154	7	ADP074421 Pseudomon
34	18	75.0	154	7	ADP074421 Pseudomon
35	18	75.0	159	8	ADP30761 Human sec
36	18	75.0	162	8	ADP30758 Human sec
37	18	75.0	162	8	ADP30756 Human sec
38	18	75.0	162	8	ADP30757 Human sec
39	18	75.0	162	8	ADP30813 Human sec
40	18	75.0	165	8	ADP31269 Human sec
41	18	75.0	168	8	ADP31099 Human sec
42	18	75.0	171	8	ADP30794 Human sec
43	18	75.0	171	8	ADP30793 Human sec
44	18	75.0	177	3	AAW10193 Arabidops
45	18	75.0	177	8	ADP30755 Human sec
46	18	75.0	177	8	ADP31682 Human sec
47	18	75.0	177	8	ADP56379 Plant pol
48	18	75.0	182	7	ABM86593 Rice abio
49	18	75.0	182	7	ABM89559 Rice abio
50	18	75.0	192	8	ADP31576 Human sec
51	18	75.0	194	7	ADP078166 Pseudomon
52	18	75.0	199	8	ADY13615 Plant ful
53	18	75.0	201	8	ADP30711 Human sec
54	18	75.0	204	3	AAW10192 Arabidops
55	18	75.0	210	8	ADP30986 Human sec
56	18	75.0	210	8	ADP30960 Human sec
57	18	75.0	219	8	ADP30593 Human sec
58	18	75.0	222	8	ADP31354 Human sec
59	18	75.0	228	8	ADP30764 Human sec
60	18	75.0	233	3	AAW74791 Neisseria
61	18	75.0	234	8	ADP31480 Human sec
62	18	75.0	234	8	ADP31481 Human sec
63	18	75.0	237	8	ADP30792 Human sec
64	18	75.0	239	7	ABO81291 Pseudomon
65	18	75.0	249	8	ADP30754 Human sec
66	18	75.0	249	8	ADP31443 Human sec
67	18	75.0	252	8	ADP31485 Human sec
68	18	75.0	258	7	ABO80689 Pseudomon
69	18	75.0	258	8	ADP30597 Human sec
70	18	75.0	267	8	ADP30822 Human sec
71	18	75.0	270	7	ABO68843 Pseudomon
72	18	75.0	276	8	ADP30568 Human sec
73	18	75.0	277	8	ADP30503 Human sec
74	18	75.0	291	8	ADP31107 Human sec
75	18	75.0	294	8	ADP31193 Human sec
76	18	75.0	294	8	ADP31045 Human sec
77	18	75.0	300	8	ADP30775 Human sec
78	18	75.0	309	8	ADP30862 Human sec
79	18	75.0	309	8	ADP30873 Human sec
80	18	75.0	310	7	ABO80230 Pseudomon
81	18	75.0	312	8	ADP31454 Human sec
82	18	75.0	318	7	ABM85364 Mouse pro
83	18	75.0	318	8	ADP30588 Human sec
84	18	75.0	318	8	ADP30824 Human sec
85	18	75.0	320	8	ADP31607 Human sec
86	18	75.0	320	8	ADP31649 Human sec
87	18	75.0	323	7	ADP074543 Pseudomon
88	18	75.0	327	7	ABM89959 Rice abio
89	18	75.0	330	4	ABR59598 Drosophi1
90	18	75.0	330	8	ADP30892 Human sec
91	18	75.0	339	8	ADP30702 Human sec
92	18	75.0	345	8	ADP31016 Human sec
93	18	75.0	345	8	ADP31683 Human sec
94	18	75.0	348	8	ADP58550 Plant pol
95	18	75.0	348	8	ADP31676 Human sec
96	18	75.0	348	8	ADP31374 Human sec
97	18	75.0	353	8	ADP31558 Human sec

98	18	75.0	354	8	ADP30783	Adp30783 Human sec	171	18	75.0	558	8	ADP31255	Adp31255 Human sec
99	18	75.0	354	8	ADP30779	Adp30779 Human sec	172	18	75.0	564	8	ADP31194	Adp31194 Human sec
100	18	75.0	354	8	ADP30779	Adp30779 Human sec	173	18	75.0	574	8	ADP31133	Adp31133 Human sec
101	18	75.0	354	8	ADP30784	Adp30784 Human sec	174	18	75.0	576	8	ADP31248	Adp31248 Human sec
102	18	75.0	355	4	ADP30778	Adp30778 Human sec	175	18	75.0	583	2	ADP31358	Adp31358 Vertebrat
103	18	75.0	355	4	ABM57739	Abm57739 Drosophila	176	18	75.0	588	8	ADP31680	Adp31680 Human sec
104	18	75.0	355	7	ABM87551	Abm87551 Rice abio	177	18	75.0	592	8	ADP30917	Adp30917 Human sec
105	18	75.0	355	9	AD276101	Ad276101 Fruit fly	178	18	75.0	594	8	ADP30895	Adp30895 Human sec
106	18	75.0	357	8	ADP31223	Adp31223 Human sec	179	18	75.0	598	8	ADP311789	Adp311789 Human the
107	18	75.0	357	8	ADP31267	Adp31267 Human sec	180	18	75.0	603	8	ADP31150	Adp31150 Human sec
108	18	75.0	360	8	ADP31439	Adp31439 Human sec	181	18	75.0	604	8	ADP30940	Adp30940 Human sec
109	18	75.0	366	8	ADP31106	Adp31106 Human sec	182	18	75.0	604	8	ADP30941	Adp30941 Human sec
110	18	75.0	366	8	ADP31670	Adp31670 Human sec	183	18	75.0	605	8	ADP30507	Adp30507 Human sec
111	18	75.0	367	5	AAO21694	Aao21694 Human sec	184	18	75.0	609	8	ADP31263	Adp31263 Human sec
112	18	75.0	369	7	ABM88560	Abm88560 Rice abio	185	18	75.0	612	8	ADP31064	Adp31064 Human sec
113	18	75.0	370	8	ADY09767	Ady09767 Plant ful	186	18	75.0	615	8	ADP31132	Adp31132 Human sec
114	18	75.0	372	8	ADP31610	Adp31610 Human sec	187	18	75.0	615	8	ADP31361	Adp31361 Human sec
115	18	75.0	382	8	ADP31216	Adp31216 Human sec	188	18	75.0	615	8	ADP31360	Adp31360 Human sec
116	18	75.0	382	8	ADP31221	Adp31221 Human sec	189	18	75.0	617	8	ADP31657	Adp31657 Human sec
117	18	75.0	387	8	ADP31158	Adp31158 Human sec	190	18	75.0	618	8	ADP31554	Adp31554 Human sec
118	18	75.0	390	8	ADP31218	Adp31218 Human sec	191	18	75.0	621	8	ADP30896	Adp30896 Human sec
119	18	75.0	393	8	ADP31345	Adp31345 Human sec	192	18	75.0	639	8	ADP31521	Adp31521 Human sec
120	18	75.0	399	8	ADP31464	Adp31464 Human sec	193	18	75.0	639	8	ADP30861	Adp30861 Human sec
121	18	75.0	401	7	ABO68344	Aboc68344 Pseudomon	194	18	75.0	642	8	ADP31265	Adp31265 Human sec
122	18	75.0	411	8	ADP31104	Adp31104 Human sec	195	18	75.0	645	8	ADP31124	Adp31124 Human sec
123	18	75.0	411	8	ADP30729	Adp30729 Human sec	196	18	75.0	645	8	ADP31125	Adp31125 Human sec
124	18	75.0	420	8	ADP31349	Adp31349 Human sec	197	18	75.0	645	8	ADP30858	Adp30858 Human sec
125	18	75.0	421	8	ADP31159	Adp31159 Human sec	198	18	75.0	645	8	ADP31183	Adp31183 Human sec
126	18	75.0	423	8	ADP31479	Adp31479 Human sec	199	18	75.0	654	8	ADP30666	Adp30666 Human sec
127	18	75.0	425	8	ADP30549	Adp30549 Human sec	200	18	75.0	658	8	ADP31226	Adp31226 Human sec
128	18	75.0	429	8	ADP31207	Adp31207 Human sec	201	18	75.0	665	8	ADP30571	Adp30571 Human sec
129	18	75.0	434	5	ABP41972	Abp41972 Human cva	202	18	75.0	666	8	ADP31547	Adp31547 Human sec
130	18	75.0	441	8	ADP31206	Adp31206 Human sec	203	18	75.0	666	8	ADP31256	Adp31256 Human sec
131	18	75.0	449	8	ADP31283	Adp31283 Human sec	204	18	75.0	669	8	ADP31598	Adp31598 Human sec
132	18	75.0	450	8	ADP30887	Adp30887 Human sec	205	18	75.0	669	8	ADP31493	Adp31493 Human sec
133	18	75.0	453	8	ADP31465	Adp31465 Human sec	206	18	75.0	677	8	ADP31585	Adp31585 Human sec
134	18	75.0	471	8	ADP31567	Adp31567 Human sec	207	18	75.0	681	8	ADP30868	Adp30868 Human sec
135	18	75.0	478	8	ADP31007	Adp31007 Human sec	208	18	75.0	681	8	ADP31053	Adp31053 Human sec
136	18	75.0	483	8	ADP30852	Adp30852 Human sec	209	18	75.0	699	8	ADP31080	Adp31080 Human sec
137	18	75.0	483	8	ADP30853	Adp30853 Human sec	210	18	75.0	699	8	ADP31079	Adp31079 Human sec
138	18	75.0	492	4	AAH71707	Aah71707 Human erf	211	18	75.0	699	8	ADP31241	Adp31241 Human sec
139	18	75.0	492	7	ADD14170	Add14170 Human src	212	18	75.0	699	8	ADP31240	Adp31240 Human sec
140	18	75.0	494	6	ABR64215	Abt64215 Angiogene	213	18	75.0	702	8	ADP31518	Adp31518 Human sec
141	18	75.0	494	8	ADO19178	Ado19178 Human PRO	214	18	75.0	711	8	ADP31215	Adp31215 Human sec
142	18	75.0	494	8	ADO19174	Ado19174 Human PRO	215	18	75.0	711	8	ADP31652	Adp31652 Human sec
143	18	75.0	494	8	ADP54221	Adp54221 Human PRO	216	18	75.0	711	8	ADP31535	Adp31535 Human sec
144	18	75.0	494	9	ADX05928	Adx05928 Cyclin-de	217	18	75.0	711	8	ADP31440	Adp31440 Human sec
145	18	75.0	494	9	ADY14502	Ady14502 PRO polyp	218	18	75.0	714	8	ADP31561	Adp31561 Human sec
146	18	75.0	494	9	ADY14498	Ady14498 PRO polyp	219	18	75.0	720	5	AAO14994	Aao14994 Lamtlnh-r
147	18	75.0	502	7	ADC35086	Adc35086 Human bre	220	18	75.0	720	9	ADP47628	Adp47628 Lamtlnh-5
148	18	75.0	513	8	ADP31681	Adp31681 Human sec	221	18	75.0	722	8	ADP98843	Adp98843 C. albica
149	18	75.0	514	8	ADP31122	Adp31122 Human sec	222	18	75.0	750	8	ADP31131	Adp31131 Human sec
150	18	75.0	516	7	ABM85365	Abm85365 Human pro	223	18	75.0	756	8	ADP31039	Adp31039 Human sec
151	18	75.0	516	8	ADP30871	Adp30871 Human sec	224	18	75.0	757	8	ADP30925	Adp30925 Human sec
152	18	75.0	516	8	ADP31418	Adp31418 Human sec	225	18	75.0	765	8	ADP31149	Adp31149 Human sec
153	18	75.0	531	8	ADP31539	Adp31539 Human sec	226	18	75.0	768	8	ADP30912	Adp30912 Human sec
154	18	75.0	531	8	ADP30594	Adp30594 Human sec	227	18	75.0	768	8	ADP31362	Adp31362 Human sec
155	18	75.0	531	8	ADP31540	Adp31540 Human sec	228	18	75.0	768	8	ADP31363	Adp31363 Human sec
156	18	75.0	531	8	ADP31696	Adp31696 Human sec	229	18	75.0	771	8	ADP30907	Adp30907 Human sec
157	18	75.0	537	8	ADP31557	Adp31557 Human sec	230	18	75.0	771	8	ADP31469	Adp31469 Human sec
158	18	75.0	538	8	ADY22481	Ady22481 Plant ful	231	18	75.0	771	8	ADP31244	Adp31244 Human sec
159	18	75.0	541	2	ADH11359	Adh11359 Vertebrat	232	18	75.0	774	8	ADP30885	Adp30885 Human sec
160	18	75.0	546	8	ADP30864	Adp30864 Human sec	233	18	75.0	774	8	ADP30506	Adp30506 Human sec
161	18	75.0	546	8	ADO19176	Ado19176 Human PRO	234	18	75.0	774	8	ADP31225	Adp31225 Human sec
162	18	75.0	546	8	ADP31449	Adp31449 Human sec	235	18	75.0	779	8	ADP30915	Adp30915 Human sec
163	18	75.0	546	8	ADP54223	Adp54223 Human PRO	236	18	75.0	779	8	ADP30897	Adp30897 Human sec
164	18	75.0	546	9	ADY14500	Ady14500 PRO polyp	237	18	75.0	780	8	ADP31077	Adp31077 Human sec
165	18	75.0	549	8	ADP31009	Adp31009 Human sec	238	18	75.0	783	8	ADP31436	Adp31436 Human sec
166	18	75.0	552	8	ADP30875	Adp30875 Human sec	239	18	75.0	783	8	ADP30496	Adp30496 Human sec
167	18	75.0	552	8	ADP31019	Adp31019 Human sec	240	18	75.0	786	8	ADP30494	Adp30494 Human sec
168	18	75.0	555	8	ADP31168	Adp31168 Human sec	241	18	75.0	795	8	ADP31332	Adp31332 Human sec
169	18	75.0	557	8	ADP31103	Adp31103 Human sec	242	18	75.0	804	8	ADP31635	Adp31635 Human sec
170	18	75.0	558	8	ADP31254	Adp31254 Human sec	243	18	75.0	804	8	ADP31291	Adp31291 Human sec

244	18	75.0	807	8	ADP31036	Adp31036 Human sec	317	18	75.0	1092	8	ADP31153	Adp31153 Human sec
245	18	75.0	810	8	ADP31307	Adp31307 Human sec	318	18	75.0	1092	8	ADP31358	Adp31358 Human sec
246	18	75.0	813	8	ADP30649	Adp30649 Human sec	319	18	75.0	1107	8	ADP30482	Adp30482 Human sec
247	18	75.0	813	8	ADP30561	Adp30561 Human sec	320	18	75.0	1107	8	ADP30504	Adp30504 Human sec
248	18	75.0	816	8	ADP31249	Adp31249 Human sec	321	18	75.0	1110	8	ADP31430	Adp31430 Human sec
249	18	75.0	816	8	ADP31581	Adp31581 Human sec	322	18	75.0	1113	8	ADP31508	Adp31508 Human sec
250	18	75.0	828	8	ADP31569	Adp31569 Human sec	323	18	75.0	1116	8	ADP31203	Adp31203 Human sec
251	18	75.0	831	8	ADP31333	Adp31333 Human sec	324	18	75.0	1116	8	ADP31692	Adp31692 Human sec
252	18	75.0	831	8	ADP31123	Adp31123 Human sec	325	18	75.0	1116	8	ADP31202	Adp31202 Human sec
253	18	75.0	831	8	ADP31179	Adp31179 Human sec	326	18	75.0	1128	6	ADA15725	C. elegans
254	18	75.0	843	3	AAy444348	AAy444348 Human hep	327	18	75.0	1133	8	ADP30931	Adp30931 Human sec
255	18	75.0	843	8	ADP30663	Adp30663 Human sec	328	18	75.0	1133	8	ADP30928	Adp30928 Human sec
256	18	75.0	849	8	ADP31182	Adp31182 Human sec	329	18	75.0	1133	8	ADP30933	Adp30933 Human sec
257	18	75.0	861	8	ADP31020	Adp31020 Human sec	330	18	75.0	1133	8	ADP30939	Adp30939 Human sec
258	18	75.0	861	8	ADP31021	Adp31021 Human sec	331	18	75.0	1134	8	ADP30647	Adp30647 Human sec
259	18	75.0	865	8	ADP30566	Adp30566 Human sec	332	18	75.0	1140	8	ADP31128	Adp31128 Human sec
260	18	75.0	869	7	ABO68000	ABO68000 Pseudomon	333	18	75.0	1140	8	ADP31130	Adp31130 Human sec
261	18	75.0	870	8	ADP30646	Adp30646 Human sec	334	18	75.0	1141	8	ADP30677	Adp30677 Human sec
262	18	75.0	882	8	ADP31688	Adp31688 Human sec	335	18	75.0	1142	8	ADP30929	Adp30929 Human sec
263	18	75.0	885	8	ADP31198	Adp31198 Human sec	336	18	75.0	1155	8	ADP31455	Adp31455 Human sec
264	18	75.0	887	8	ADP30554	Adp30554 Human sec	337	18	75.0	1165	8	ADP30879	Adp30879 Human sec
265	18	75.0	887	8	ADP30548	Adp30548 Human sec	338	18	75.0	1168	8	ADP31046	Adp31046 Human sec
266	18	75.0	889	8	ADP31648	Adp31648 Human sec	339	18	75.0	1170	8	ADP30919	Adp30919 Human sec
267	18	75.0	890	8	ADP31059	Adp31059 Human sec	340	18	75.0	1170	8	ADP30922	Adp30922 Human sec
268	18	75.0	892	8	ADP31578	Adp31578 Human sec	341	18	75.0	1173	8	ADP31155	Adp31155 Human sec
269	18	75.0	897	8	ADP30914	Adp30914 Human sec	342	18	75.0	1189	8	ADP31043	Adp31043 Human sec
270	18	75.0	906	8	ADP31409	Adp31409 Human sec	343	18	75.0	1192	8	ADP30964	Adp30964 Human sec
271	18	75.0	918	8	ADP31459	Adp31459 Human sec	344	18	75.0	1192	8	ADP31180	Adp31180 Human sec
272	18	75.0	921	8	ADP31522	Adp31522 Human sec	345	18	75.0	1197	8	ADP31034	Adp31034 Human sec
273	18	75.0	925	5	AAO14246	AAO14246 Human pre	346	18	75.0	1199	8	ADP31044	Adp31044 Human sec
274	18	75.0	930	8	ADP31444	Adp31444 Human sec	347	18	75.0	1215	8	ADP30898	Adp30898 Human sec
275	18	75.0	933	8	ADP31510	Adp31510 Human sec	348	18	75.0	1215	8	ADP30900	Adp30900 Human sec
276	18	75.0	936	8	ADP31597	Adp31597 Human sec	349	18	75.0	1221	8	ADP31293	Adp31293 Human sec
277	18	75.0	936	8	ADP31568	Adp31568 Human sec	350	18	75.0	1224	8	ADP31426	Adp31426 Human sec
278	18	75.0	939	8	ADP31542	Adp31542 Human sec	351	18	75.0	1233	8	ADP30523	Adp30523 Human sec
279	18	75.0	939	8	ADP31086	Adp31086 Human sec	352	18	75.0	1239	8	ADP31297	Adp31297 Human sec
280	18	75.0	939	8	ADP31541	Adp31541 Human sec	353	18	75.0	1252	8	ADP30678	Adp30678 Human sec
281	18	75.0	945	8	ADP31238	Adp31238 Human sec	354	18	75.0	1260	8	ADP31533	Adp31533 Human sec
282	18	75.0	945	8	ADP31237	Adp31237 Human sec	355	18	75.0	1269	8	ADP31382	Adp31382 Human sec
283	18	75.0	947	8	ADP30937	Adp30937 Human sec	356	18	75.0	1269	8	ADP31381	Adp31381 Human sec
284	18	75.0	948	8	ADP30586	Adp30586 Human sec	357	18	75.0	1282	8	ADP31328	Adp31328 Human sec
285	18	75.0	950	8	ADP31167	Adp31167 Human sec	358	18	75.0	1288	8	ADP31114	Adp31114 Human sec
286	18	75.0	954	8	ADP31015	Adp31015 Human sec	359	18	75.0	1289	8	ADP30675	Adp30675 Human sec
287	18	75.0	957	8	ADP31528	Adp31528 Human sec	360	18	75.0	1302	8	ADP31695	Adp31695 Human sec
288	18	75.0	960	8	ADP31471	Adp31471 Human sec	361	18	75.0	1312	8	ADP30999	Adp30999 Human sec
289	18	75.0	960	8	ADP31470	Adp31470 Human sec	362	18	75.0	1358	8	ADP30995	Adp30995 Human sec
290	18	75.0	966	8	ADP30745	Adp30745 Human sec	363	18	75.0	1359	8	ADP31129	Adp31129 Human sec
291	18	75.0	975	8	ADP30843	Adp30843 Human sec	364	18	75.0	1362	8	ADP31181	Adp31181 Human sec
292	18	75.0	990	8	ADP31553	Adp31553 Human sec	365	18	75.0	1365	8	ADP31035	Adp31035 Human sec
293	18	75.0	1002	8	ADG39639	Adg39639 Human pan	366	18	75.0	1371	8	ADP31646	Adp31646 Human sec
294	18	75.0	1002	8	ADP30866	Adp30866 Human sec	367	18	75.0	1371	8	ADP30876	Adp30876 Human sec
295	18	75.0	1010	8	ADP31296	Adp31296 Human sec	368	18	75.0	1380	8	ADP31566	Adp31566 Human sec
296	18	75.0	1017	6	AAO26719	AAO26719 SR proteol	369	18	75.0	1383	8	ADP31091	Adp31091 Human sec
297	18	75.0	1017	8	ADP31268	Adp31268 Human sec	370	18	75.0	1387	8	ADP30946	Adp30946 Human sec
298	18	75.0	1023	8	ADP30994	Adp30994 Human sec	371	18	75.0	1404	8	ADP31627	Adp31627 Human sec
299	18	75.0	1030	8	ADP30913	Adp30913 Human sec	372	18	75.0	1420	8	ADP30944	Adp30944 Human sec
300	18	75.0	1033	8	ADP30984	Adp30984 Human sec	373	18	75.0	1431	8	ADP31609	Adp31609 Human sec
301	18	75.0	1035	8	ADP31552	Adp31552 Human sec	374	18	75.0	1437	8	ADP31357	Adp31357 Human sec
302	18	75.0	1038	8	ADP30860	Adp30860 Human sec	375	18	75.0	1440	8	ADG34533	Glucocort
303	18	75.0	1048	8	ADP31642	Adp31642 Human sec	376	18	75.0	1454	8	ADP31177	Adp31177 Human sec
304	18	75.0	1056	8	ADP31082	Adp31082 Human sec	377	18	75.0	1456	8	ADP30923	Adp30923 Human sec
305	18	75.0	1057	8	ADP31592	Adp31592 Human sec	378	18	75.0	1464	8	ADP31040	Adp31040 Human sec
306	18	75.0	1059	8	ADP31042	Adp31042 Human sec	379	18	75.0	1464	8	ADP31437	Adp31437 Human sec
307	18	75.0	1065	8	ADP30973	Adp30973 Human sec	380	18	75.0	1470	8	ADP31290	Adp31290 Human sec
308	18	75.0	1065	8	ADP31347	Adp31347 Human sec	381	18	75.0	1472	8	ADP31611	Adp31611 Human sec
309	18	75.0	1076	6	ABU81145	ABU81145 Human PRO	382	18	75.0	1480	8	ADP30557	Adp30557 Human sec
310	18	75.0	1076	6	ABU66845	ABU66845 Human PRO	383	18	75.0	1485	8	ADP31383	Adp31383 Human sec
311	18	75.0	1081	2	AAy24319	AAy24319 Mouse dep	384	18	75.0	1485	8	ADP31384	Adp31384 Human sec
312	18	75.0	1083	8	ADP31073	Adp31073 Human sec	385	18	75.0	1488	8	ADP31385	Adp31385 Human sec
313	18	75.0	1086	8	ADP31175	Adp31175 Human sec	386	18	75.0	1488	8	ADP31386	Adp31386 Human sec
314	18	75.0	1086	8	ADP31447	Adp31447 Human sec	387	18	75.0	1498	8	ADP30684	Adp30684 Human sec
315	18	75.0	1086	8	ADP31629	Adp31629 Human sec	388	18	75.0	1506	8	ADP30596	Adp30596 Human sec
316	18	75.0	1090	8	ADQ10187	ADQ10187 Human pol	389	18	75.0	1510	7	ADG30698	Xanthomon

390	18	75.0	1518	8	ADP31532	Adp31532	Human	sec	463	18	75.0	1894	6	ABO05229	AbO05229	Novel	hum
391	18	75.0	1530	8	ADP31536	Adp31536	Human	sec	464	18	75.0	1917	6	ADAI5719	AdAI5719	C. elegan	
392	18	75.0	1539	8	ADP31201	Adp31201	Human	sec	465	18	75.0	1933	8	ADP30889	Adp30889	Human	sec
393	18	75.0	1539	8	ADP31200	Adp31200	Human	sec	466	18	75.0	1933	8	ADP30902	Adp30902	Human	sec
394	18	75.0	1550	8	ADP30567	Adp30567	Human	sec	467	18	75.0	2001	8	ADP31644	Adp31644	Human	sec
395	18	75.0	1584	8	ADP31405	Adp31405	Human	sec	468	18	75.0	2020	8	ADP31056	Adp31056	Human	sec
396	18	75.0	1587	8	ADP30591	Adp30591	Human	sec	469	18	75.0	2020	8	ADP30511	Adp30511	Human	sec
397	18	75.0	1614	8	ADP31529	Adp31529	Human	sec	470	18	75.0	2020	8	ADP30512	Adp30512	Human	sec
398	18	75.0	1617	8	ADP30660	Adp30660	Human	sec	471	18	75.0	2020	8	ADP30509	Adp30509	Human	sec
399	18	75.0	1623	8	ADP30552	Adp30552	Human	sec	472	18	75.0	2027	8	ADP31058	Adp31058	Human	sec
400	18	75.0	1629	8	ADP30945	Adp30945	Human	sec	473	18	75.0	2044	8	ADP30510	Adp30510	Human	sec
401	18	75.0	1647	8	ADP30670	Adp30670	Human	sec	474	18	75.0	2088	8	ADP31178	Adp31178	Human	sec
402	18	75.0	1647	8	ADP31052	Adp31052	Human	sec	475	18	75.0	2091	8	ADP31088	Adp31088	Human	sec
403	18	75.0	1652	6	ADAI5715	AdAI5715	C. elegan		476	18	75.0	2123	8	ADP30657	Adp30657	Human	sec
404	18	75.0	1665	8	ADP31187	Adp31187	Human	sec	477	18	75.0	2123	8	ADP30564	Adp30564	Human	sec
405	18	75.0	1719	8	ADP31137	Adp31137	Human	sec	478	18	75.0	2127	8	ADP31327	Adp31327	Human	sec
406	18	75.0	1725	8	ADP30654	Adp30654	Human	sec	479	18	75.0	2148	8	ADP30974	Adp30974	Human	sec
407	18	75.0	1743	6	ABU88255	AbU88255	Novel	hum	480	18	75.0	2187	8	ADP30882	Adp30882	Human	sec
408	18	75.0	1743	6	ABU90134	AbU90134	Novel	hum	481	18	75.0	2260	8	ADP30687	Adp30687	Human	sec
409	18	75.0	1743	6	ABU96436	AbU96436	Novel	hum	482	18	75.0	2272	8	ADP30669	Adp30669	Human	sec
410	18	75.0	1743	6	ABU99045	AbU99045	Novel	hum	483	18	75.0	2304	8	ADP31252	Adp31252	Human	sec
411	18	75.0	1743	6	ABU98260	AbU98260	Novel	hum	484	18	75.0	2307	8	ADP31394	Adp31394	Human	sec
412	18	75.0	1743	6	ABU91966	AbU91966	Novel	hum	485	18	75.0	2349	8	ADP30959	Adp30959	Human	sec
413	18	75.0	1743	6	ABU85270	AbU85270	Novel	hum	486	18	75.0	2358	8	ADP31690	Adp31690	Human	sec
414	18	75.0	1743	6	ABO00409	AbO00409	Novel	hum	487	18	75.0	2382	8	ADP31341	Adp31341	Human	sec
415	18	75.0	1743	6	ABU88960	AbU88960	Novel	hum	488	18	75.0	2391	8	ADP31366	Adp31366	Human	sec
416	18	75.0	1743	6	ABO06456	AbO06456	Novel	hum	489	18	75.0	2415	8	ADP31023	Adp31023	Human	sec
417	18	75.0	1743	6	ABU95516	AbU95516	Novel	hum	490	18	75.0	2418	8	ADP31105	Adp31105	Human	sec
418	18	75.0	1743	6	ABU95206	AbU95206	Novel	hum	491	18	75.0	2454	8	ADP30469	Adp30469	Human	sec
419	18	75.0	1743	6	ABU90754	AbU90754	Novel	hum	492	18	75.0	2468	8	ADP30947	Adp30947	Human	sec
420	18	75.0	1743	6	ABU93916	AbU93916	Novel	hum	493	18	75.0	2469	8	ADP31031	Adp31031	Human	sec
421	18	75.0	1743	6	ABU86190	AbU86190	Novel	hum	494	18	75.0	2484	8	ADP66690	Adp66690	Human	sec
422	18	75.0	1743	6	ABU82045	AbU82045	Novel	hum	495	18	75.0	2508	6	ADAI5721	AdAI5721	C. elegan	
423	18	75.0	1743	6	ABU07906	AbU07906	Novel	hum	496	18	75.0	2514	8	ADP30727	Adp30727	Human	sec
424	18	75.0	1743	6	ABU94226	AbU94226	Novel	hum	497	18	75.0	2535	8	ADP31146	Adp31146	Human	sec
425	18	75.0	1743	6	ABO00099	AbO00099	Novel	hum	498	18	75.0	2544	6	ADAI5717	AdAI5717	C. elegan	
426	18	75.0	1743	6	ABU87110	AbU87110	Novel	hum	499	18	75.0	2601	6	ADAI5723	AdAI5723	C. elegan	
427	18	75.0	1743	6	ABU91351	AbU91351	Novel	hum	500	18	75.0	2611	8	ADP31027	Adp31027	Human	sec
428	18	75.0	1743	6	ABU90444	AbU90444	Novel	hum	501	18	75.0	2616	8	ADP31253	Adp31253	Human	sec
429	18	75.0	1743	6	ABU97035	AbU97035	Novel	hum	502	18	75.0	2616	9	ABE49675	AbE49675	N. mening	
430	18	75.0	1743	6	ABO05231	AbO05231	Novel	hum	503	18	75.0	2664	8	ADP31546	Adp31546	Human	sec
431	18	75.0	1755	8	ADP31446	Adp31446	Human	sec	504	18	75.0	2688	8	ADN11593	Adn11593	Human	CD9
432	18	75.0	1766	9	ADP62715	Adp62715	Human	alp	505	18	75.0	2808	8	ADP31258	Adp31258	Human	sec
433	18	75.0	1782	8	ADP31391	Adp31391	Human	sec	506	18	75.0	2832	8	ADP31121	Adp31121	Human	sec
434	18	75.0	1782	8	ADP31270	Adp31270	Human	sec	507	18	75.0	2835	8	ADP30572	Adp30572	Human	sec
435	18	75.0	1789	8	ADP30962	Adp30962	Human	sec	508	18	75.0	2943	8	ADP31120	Adp31120	Human	sec
436	18	75.0	1803	8	ADP30562	Adp30562	Human	sec	509	18	75.0	2980	9	AEA36049	AEa36049	Maize	Sta
437	18	75.0	1827	8	ADP31170	Adp31170	Human	sec	510	18	75.0	3046	8	ADQ10419	Adq10419	Human	pol
438	18	75.0	1833	8	ADP30642	Adp30642	Human	sec	511	18	75.0	3070	8	ADP30710	Adp30710	Human	sec
439	18	75.0	1848	6	ADP31372	Adp31372	Human	sec	512	18	75.0	3070	8	ADP30700	Adp30700	Human	sec
440	18	75.0	1894	6	ABU88253	AbU88253	Novel	hum	513	18	75.0	3144	8	ADP31544	Adp31544	Human	sec
441	18	75.0	1894	6	ABU90132	AbU90132	Novel	hum	514	18	75.0	3201	8	ADP31545	Adp31545	Human	sec
442	18	75.0	1894	6	ABU96634	AbU96634	Novel	hum	515	18	75.0	3316	8	ADP31116	Adp31116	Human	sec
443	18	75.0	1894	6	ABU99043	AbU99043	Novel	hum	516	18	75.0	3339	8	ADP31219	Adp31219	Human	sec
444	18	75.0	1894	6	ABU98258	AbU98258	Novel	hum	517	18	75.0	3390	9	ABE87634	AbE87634	Human	ino
445	18	75.0	1894	6	ABU91964	AbU91964	Novel	hum	518	18	75.0	3398	8	ADP31148	Adp31148	Human	sec
446	18	75.0	1894	6	ABU85268	AbU85268	Novel	hum	519	18	75.0	3411	8	ADP30667	Adp30667	Human	sec
447	18	75.0	1894	6	ABO00407	AbO00407	Novel	hum	520	18	75.0	3447	8	ADP31112	Adp31112	Human	sec
448	18	75.0	1894	6	ABU88958	AbU88958	Novel	hum	521	18	75.0	3465	8	ADP31234	Adp31234	Human	sec
449	18	75.0	1894	6	ABO06454	AbO06454	Novel	hum	522	18	75.0	3477	8	ADP30704	Adp30704	Human	sec
450	18	75.0	1894	6	ABU95514	AbU95514	Novel	hum	523	18	75.0	3579	8	ADP31098	Adp31098	Human	sec
451	18	75.0	1894	6	ABU95204	AbU95204	Novel	hum	524	18	75.0	3585	8	ADP31117	Adp31117	Human	sec
452	18	75.0	1894	6	ABU90752	AbU90752	Novel	hum	525	18	75.0	4440	6	ABU88256	AbU88256	Novel	hum
453	18	75.0	1894	6	ABU93914	AbU93914	Novel	hum	526	18	75.0	4440	6	ABU90135	AbU90135	Novel	hum
454	18	75.0	1894	6	ABU86188	AbU86188	Novel	hum	527	18	75.0	4440	6	ABU96437	AbU96437	Novel	hum
455	18	75.0	1894	6	ABU82043	AbU82043	Novel	hum	528	18	75.0	4440	6	ABU99046	AbU99046	Novel	hum
456	18	75.0	1894	6	ABU07904	AbU07904	Novel	hum	529	18	75.0	4440	6	ABU98261	AbU98261	Novel	hum
457	18	75.0	1894	6	ABU94224	AbU94224	Novel	hum	530	18	75.0	4440	6	ABU91967	AbU91967	Novel	hum
458	18	75.0	1894	6	ABO00097	AbO00097	Novel	hum	531	18	75.0	4440	6	ABU85271	AbU85271	Novel	hum
459	18	75.0	1894	6	ABU87108	AbU87108	Novel	hum	532	18	75.0	4440	6	ABO00410	AbO00410	Novel	hum
460	18	75.0	1894	6	ABU91349	AbU91349	Novel	hum	533	18	75.0	4440	6	ABU88961	AbU88961	Novel	hum
461	18	75.0	1894	6	ABU90442	AbU90442	Novel	hum	534	18	75.0	4440	6	ABO06457	AbO06457	Novel	hum
462	18	75.0	1894	6	ABU97033	AbU97033	Novel	hum	535	18	75.0	4440	6	ABU95517	AbU95517	Novel	hum

536	18	75.0	4440	6	ABU95207	Novel	hum	609	17	70.8	53	6	ABM54353	Abm54353	Propionib
537	18	75.0	4440	6	ABU90755	Novel	hum	610	17	70.8	53	6	ABM53320	Abm53320	Propionib
538	18	75.0	4440	6	ABU93917	Novel	hum	611	17	70.8	54	6	AEA30185	Aea30185	Pertussis
539	18	75.0	4440	6	ABU86191	Novel	hum	612	17	70.8	58	4	AAU54758	AAU54758	Propionib
540	18	75.0	4440	6	ABU82046	Novel	hum	613	17	70.8	58	6	ABM51277	Abm51277	Propionib
541	18	75.0	4440	6	ABU07907	Novel	hum	614	17	70.8	58	9	ADZ11761	Adz11761	E. coli D
542	18	75.0	4440	6	ABU94227	Novel	hum	615	17	70.8	58	9	AEA30433	Aea30433	Pertussis
543	18	75.0	4440	6	ABO00100	Novel	hum	616	17	70.8	58	9	AEA30308	Aea30308	Pertussis
544	18	75.0	4440	6	ABU87111	Novel	hum	617	17	70.8	59	4	AAU51154	AAU51154	Propionib
545	18	75.0	4440	6	ABU91352	Novel	hum	618	17	70.8	59	6	ABM47673	Abm47673	Propionib
546	18	75.0	4440	6	ABU90445	Novel	hum	619	17	70.8	60	9	AEA30430	Aea30430	Pertussis
547	18	75.0	4440	6	ABU97036	Novel	hum	620	17	70.8	60	9	AEA30259	Aea30259	Pertussis
548	18	75.0	4440	6	ABO05232	Novel	hum	621	17	70.8	60	9	AEA30493	Aea30493	Pertussis
549	18	75.0	4683	8	ADP31260	Human	sec	622	17	70.8	60	9	AEA30376	Aea30376	Pertussis
550	18	75.0	4752	8	ADP30585	Human	sec	623	17	70.8	60	9	AEA30494	Aea30494	Pertussis
551	18	75.0	4752	8	ADP30651	Human	sec	624	17	70.8	60	9	AEA30431	Aea30431	Pertussis
552	18	75.0	4752	8	ADP31259	Human	sec	625	17	70.8	60	9	AEA30432	Aea30432	Pertussis
553	18	75.0	4804	8	ADP31259	Human	sec	626	17	70.8	62	7	ABO69433	ABO69433	Pseudomon
554	18	75.0	5397	8	ADP31068	Human	sec	627	17	70.8	62	7	AAV12328	AAV12328	Human 5'
555	18	75.0	5514	8	ADP31186	Human	sec	628	17	70.8	64	2	AAW67942	AAW67942	Fragment
556	18	75.0	5514	8	ADP31591	Human	sec	629	17	70.8	64	4	AAW85119	AAW85119	Human imm
557	18	75.0	5820	8	ADP31118	Human	sec	630	17	70.8	64	5	ABP69086	ABP69086	Human pol
558	18	75.0	6465	8	ADP30705	Human	sec	631	17	70.8	64	5	ABG13638	ABG13638	Novel
559	18	75.0	6729	8	ABJ38280	Human	sec	632	17	70.8	66	4	ABM86677	ABM86677	Rice
560	18	75.0	7285	6	AAO16358	Human	tra	633	17	70.8	68	7	ABO78891	ABO78891	Human ORF
561	18	75.0	7339	6	ADP31119	Human	sec	634	17	70.8	68	5	ABP09377	ABP09377	Human ORF
562	18	75.0	8973	8	ADP31425	Human	sec	635	17	70.8	70	4	AAU53345	AAU53345	Propionib
563	18	75.0	8976	8	ADP31425	Human	sec	636	17	70.8	70	6	ABM49864	ABM49864	Propionib
564	18	75.0	9195	8	ADP31494	Human	sec	637	17	70.8	71	2	AAW83934	AAW83934	Human sec
565	18	75.0	10944	8	ADP31311	Human	sec	638	17	70.8	71	3	AAW834715	AAW834715	Human sec
566	18	75.0	11328	8	ADP31310	Human	sec	639	17	70.8	73	7	ABO82698	ABO82698	Pseudomon
567	17	70.8	8	5	ABR47073	Desmocoll		640	17	70.8	73	8	ADP29540	ADP29540	Human sec
568	17	70.8	8	5	ABR46610	Desmocoll		641	17	70.8	74	4	AAU59878	AAU59878	Propionib
569	17	70.8	8	5	ADZ50990	Amino aci		642	17	70.8	74	6	ABM56397	ABM56397	Propionib
570	17	70.8	8	9	ADZ50974	Amino aci		643	17	70.8	74	6	ADV76639	ADV76639	Human TBR
571	17	70.8	10	4	AAW5975	Human	com	644	17	70.8	77	9	ADW73420	ADW73420	Plant ful
572	17	70.8	10	4	AAW5917	Human	com	645	17	70.8	78	9	AAW94485	AAW94485	Human
573	17	70.8	20	2	AAW23344	Peptide s		646	17	70.8	79	4	AAW19029	AAW19029	Peptide #
574	17	70.8	20	3	AAE19697	O-Methylt		647	17	70.8	79	4	ABR38235	ABR38235	Peptide #
575	17	70.8	20	3	AAE05831	Liquidamb		648	17	70.8	79	4	AAW31668	AAW31668	Peptide #
576	17	70.8	20	9	ADZ14872	Liquidamb		649	17	70.8	79	4	ABR23413	ABR23413	Protein
577	17	70.8	20	9	AAW33651	Peptide #		650	17	70.8	79	4	ABG53088	ABG53088	Human liv
578	17	70.8	22	4	AAW73453	Human bon		651	17	70.8	79	5	ABG41186	ABG41186	Human pep
579	17	70.8	22	4	AAW60779	Human bon		652	17	70.8	80	3	AAW11389	AAW11389	Arbidops
580	17	70.8	22	4	ABG55176	Human liv		653	17	70.8	83	4	AAU64692	AAU64692	Propionib
581	17	70.8	22	5	ABG43312	Human pep		654	17	70.8	83	6	ABM61211	ABM61211	Propionib
582	17	70.8	24	8	ADP73753	Influenza		655	17	70.8	84	3	AAW60503	AAW60503	Arbidops
583	17	70.8	28	9	AAW98776	Human cel		656	17	70.8	84	3	AAW61337	AAW61337	Propionib
584	17	70.8	28	9	ADZ47463	Human HAS		657	17	70.8	84	6	ABM57856	ABM57856	Propionib
585	17	70.8	34	5	AAU87022	Adenoviru		658	17	70.8	84	6	ABR41187	ABR41187	L. pneumo
586	17	70.8	35	6	ABU61319	Human A d		659	17	70.8	85	9	AAU51874	AAU51874	Propionib
587	17	70.8	35	6	ABU61408	Low densi		660	17	70.8	86	4	ABM48393	ABM48393	Propionib
588	17	70.8	35	8	ADP21541	Human LDL		661	17	70.8	86	6	AAU63689	AAU63689	Propionib
589	17	70.8	37	4	AAO09823	Human pol		662	17	70.8	87	4	ABM60208	ABM60208	Propionib
590	17	70.8	39	9	ABR43285	Human LRP		663	17	70.8	87	6	ADP17451	ADP17451	Mouse IL-
591	17	70.8	39	9	AAU59877	Propionib		664	17	70.8	87	7	AAU21754	AAU21754	Novel
592	17	70.8	50	6	ABM56396	Propionib		665	17	70.8	88	4	AAU45755	AAU45755	Novel
593	17	70.8	51	4	ABR39045	Peptide #		666	17	70.8	88	4	AAU20119	AAU20119	Human DNA
594	17	70.8	51	4	AAW72275	Human bon		667	17	70.8	88	5	ABG91368	ABG91368	Novel
595	17	70.8	51	4	AAW59697	Human bra		668	17	70.8	88	5	ABM42274	ABM42274	Human neo
596	17	70.8	51	4	ABG53961	Human liv		669	17	70.8	88	7	ABO69228	ABO69228	Pseudomon
597	17	70.8	51	5	ABG42090	Human pep		670	17	70.8	88	7	ADP31655	ADP31655	Human sec
598	17	70.8	53	4	ABR39811	Peptide #		671	17	70.8	90	8	AAV12296	AAV12296	Human 5'
599	17	70.8	53	4	AAW3401	Peptide #		672	17	70.8	91	2	ABO09944	ABO09944	Guinea pi
600	17	70.8	53	4	AAW3401	Peptide #		673	17	70.8	92	5	ABM09943	ABM09943	Human int
601	17	70.8	53	4	AAW3401	Peptide #		674	17	70.8	92	5	AAE39395	AAE39395	Guinea pi
602	17	70.8	53	4	AAW3401	Peptide #		675	17	70.8	92	7	AAE39395	AAE39395	Human int
603	17	70.8	53	4	AAW3401	Peptide #		676	17	70.8	92	7	AAE39395	AAE39395	Human int
604	17	70.8	53	4	AAW3401	Peptide #		677	17	70.8	92	7	AAE39395	AAE39395	Human int
605	17	70.8	53	4	AAW3401	Peptide #		678	17	70.8	92	7	AAE39395	AAE39395	Human int
606	17	70.8	53	4	AAW3401	Peptide #		679	17	70.8	92	9	AAE39395	AAE39395	Human int
607	17	70.8	53	4	AAW3401	Peptide #		680	17	70.8	92	9	AAE39395	AAE39395	Human int
608	17	70.8	53	5	ABG43038	Human pep		681	17	70.8	93	4	AAU56711	AAU56711	Propionib

682	17	70.8	93	6	ABM53230	755	17	70.8	134	6	ABM55230	Adm55230	Propionib
683	17	70.8	96	6	AAU52513	756	17	70.8	134	7	ABO76893	ABO76893	Pseudomon
684	17	70.8	96	6	ABM49032	757	17	70.8	135	3	AAV75297	AAV75297	Neisseria
685	17	70.8	97	4	AAU66144	758	17	70.8	135	3	AAV75296	AAV75296	Neisseria
686	17	70.8	97	5	ABP09938	759	17	70.8	135	5	ABP08959	ABP08959	Human ORF
687	17	70.8	97	6	ABM62663	760	17	70.8	135	7	ABO74109	ABO74109	Pseudomon
688	17	70.8	97	7	ABO79445	761	17	70.8	135	8	ADP31638	ADP31638	Human sec
689	17	70.8	97	9	ABE37882	762	17	70.8	135	8	ADP07492	ADP07492	Plant ful
690	17	70.8	99	4	AAU49262	763	17	70.8	136	3	AAV75295	AAV75295	Neisseria
691	17	70.8	99	5	ABP43926	764	17	70.8	136	7	ADM57545	ADM57545	Human his
692	17	70.8	99	6	ABM45781	765	17	70.8	136	7	ABO73988	ABO73988	Pseudomon
693	17	70.8	99	8	ADM45780	766	17	70.8	136	7	ABM66887	ABM66887	Rice abio
694	17	70.8	102	7	ABO79727	767	17	70.8	137	3	AAV65413	AAV65413	Human 5'
695	17	70.8	105	7	ADBE2627	768	17	70.8	137	8	ADU72977	ADU72977	Signal pe
696	17	70.8	105	4	AAAG64065	769	17	70.8	137	9	ADZ73968	ADZ73968	Human com
697	17	70.8	106	4	AAAG64068	770	17	70.8	140	7	ADFI3949	ADFI3949	Human end
698	17	70.8	106	4	ABBI5333	771	17	70.8	140	7	ABO82130	ABO82130	Pseudomon
699	17	70.8	106	5	AAU96154	772	17	70.8	141	7	ABO77083	ABO77083	Pseudomon
700	17	70.8	106	6	ADA57463	773	17	70.8	143	5	AAU96140	AAU96140	Human BRP
701	17	70.8	106	6	ADA41343	774	17	70.8	144	7	ABO79215	ABO79215	Pseudomon
702	17	70.8	106	6	ABR48065	775	17	70.8	144	8	ADI39724	ADI39724	TADG-14 p
703	17	70.8	106	6	ABP72423	776	17	70.8	144	8	ADI37148	ADI37148	Human pro
704	17	70.8	106	8	ABP72484	777	17	70.8	144	8	ADP31474	ADP31474	Human sec
705	17	70.8	106	9	ADM86161	778	17	70.8	144	8	ADFI3954	ADFI3954	Human end
706	17	70.8	107	7	ADFI3950	779	17	70.8	148	7	ABO74263	ABO74263	Human gen
707	17	70.8	110	8	ADP30723	780	17	70.8	149	4	AAU68926	AAU68926	Human pro
708	17	70.8	111	4	ABG03591	781	17	70.8	149	7	ADE31059	ADE31059	Human dia
709	17	70.8	112	7	ABO77337	782	17	70.8	149	7	ADE13246	ADE13246	Protease
710	17	70.8	113	6	AAU50472	783	17	70.8	149	8	ADH78455	ADH78455	Human pro
711	17	70.8	113	6	ABM46991	784	17	70.8	150	8	ADX66524	ADX66524	Plant ful
712	17	70.8	113	8	ADP31617	785	17	70.8	154	7	ABO74102	ABO74102	Pseudomon
713	17	70.8	113	8	ADP31613	786	17	70.8	154	8	ABO60522	ABO60522	Human gen
714	17	70.8	117	3	AAE25477	787	17	70.8	156	7	ABO80146	ABO80146	Pseudomon
715	17	70.8	117	7	ADBE5158	788	17	70.8	156	8	ADP30759	ADP30759	Human sec
716	17	70.8	117	7	ABO68458	789	17	70.8	157	8	ADX78826	ADX78826	Plant ful
717	17	70.8	118	6	ABU99121	790	17	70.8	162	3	AAI95932	AAI95932	Porcine a
718	17	70.8	118	8	ADM93807	791	17	70.8	162	4	AAU42696	AAU42696	Pseudomon
719	17	70.8	120	7	ADBE4538	792	17	70.8	162	6	ABM39215	ABM39215	Propionib
720	17	70.8	120	8	ADP30556	793	17	70.8	163	7	ABO80365	ABO80365	Pseudomon
721	17	70.8	122	4	AAU42386	794	17	70.8	163	7	ABO68705	ABO68705	Pseudomon
722	17	70.8	122	6	ABM38905	795	17	70.8	164	7	ABO78378	ABO78378	Pseudomon
723	17	70.8	123	7	ABO66168	796	17	70.8	165	4	AAU58075	AAU58075	Propionib
724	17	70.8	124	6	AAU51146	797	17	70.8	165	6	ABM54594	ABM54594	Propionib
725	17	70.8	124	6	ABM47665	798	17	70.8	165	8	ADP31286	ADP31286	Human sec
726	17	70.8	127	8	ADP31622	799	17	70.8	165	8	ADP31515	ADP31515	Human sec
727	17	70.8	129	4	AAAG64067	800	17	70.8	165	8	ADP31174	ADP31174	Human sec
728	17	70.8	129	5	AAU96157	801	17	70.8	165	8	ADP31516	ADP31516	Human sec
729	17	70.8	129	6	AAU96129	802	17	70.8	165	8	ADP31041	ADP31041	Human sec
730	17	70.8	129	6	ABU97054	803	17	70.8	168	8	ADP31621	ADP31621	Human sec
731	17	70.8	130	4	AAAG63211	804	17	70.8	168	8	ADP30797	ADP30797	Human sec
732	17	70.8	130	4	AAAG64064	805	17	70.8	168	8	ADP30769	ADP30769	Human sec
733	17	70.8	130	4	AAU41366	806	17	70.8	168	8	ADP30771	ADP30771	Human sec
734	17	70.8	130	4	AAE09440	807	17	70.8	171	8	ADP30599	ADP30599	Human sec
735	17	70.8	130	5	AAU97612	808	17	70.8	175	3	AAAG54788	AAAG54788	Arbidops
736	17	70.8	130	5	AAU96153	809	17	70.8	175	7	ABO80411	ABO80411	Pseudomon
737	17	70.8	130	5	AAU10368	810	17	70.8	175	7	ABO69379	ABO69379	Pseudomon
738	17	70.8	130	5	AAU10366	811	17	70.8	177	7	ABO70036	ABO70036	Pseudomon
739	17	70.8	130	5	ABG74215	812	17	70.8	178	7	ABO74750	ABO74750	Pseudomon
740	17	70.8	130	6	ABG74206	813	17	70.8	178	7	ABO73284	ABO73284	Pseudomon
741	17	70.8	130	6	ABM37885	814	17	70.8	179	7	ABO75778	ABO75778	Pseudomon
742	17	70.8	130	6	ABP72422	815	17	70.8	180	8	ADP30820	ADP30820	Human sec
743	17	70.8	130	7	ABO81121	816	17	70.8	180	8	ADP30825	ADP30825	Human sec
744	17	70.8	130	7	ABO81191	817	17	70.8	180	8	ADP30828	ADP30828	Human sec
745	17	70.8	130	8	ADP72483	818	17	70.8	180	8	ADP30821	ADP30821	Human sec
746	17	70.8	130	8	ADM86155	819	17	70.8	180	8	ADP30826	ADP30826	Human sec
747	17	70.8	130	9	ADM86175	820	17	70.8	180	8	ADP30827	ADP30827	Human sec
748	17	70.8	130	9	ABE12349	821	17	70.8	182	7	ABO78099	ABO78099	Pseudomon
749	17	70.8	131	5	AAU96155	822	17	70.8	182	7	ABO71666	ABO71666	Pseudomon
750	17	70.8	133	4	AAU49190	823	17	70.8	182	8	ADN99530	ADN99530	Novel hum
751	17	70.8	133	4	ABG04654	824	17	70.8	184	7	ABO77341	ABO77341	Pseudomon
752	17	70.8	133	6	ABM45709	825	17	70.8	185	8	ADP31109	ADP31109	Human sec
753	17	70.8	133	7	ABO82412	826	17	70.8	186	4	ABG08184	ABG08184	Novel hum
754	17	70.8	134	4	AAU58711	827	17	70.8	187	7	ABO80088	ABO80088	Pseudomon



828	17	70.8	190	6	ABM65810	901	17	70.8	249	3	AA60495	Agg60495	Arabidops
829	17	70.8	191	7	ABO76619	902	17	70.8	249	3	ABO83546	ABO83546	Pseudomon
830	17	70.8	192	8	ADP31335	903	17	70.8	250	6	ABP56993	ABP56993	E. tenell
831	17	70.8	192	8	ADP30575	904	17	70.8	250	7	ABO82689	ABO82689	Pseudomon
832	17	70.8	198	4	AAU47307	905	17	70.8	251	7	ADJ92161	ADJ92161	Human hai
833	17	70.8	198	6	ABM43826	906	17	70.8	251	7	ABO73872	ABO73872	Pseudomon
834	17	70.8	198	8	ADP30810	907	17	70.8	251	9	ABM91192	ABM91192	Human hai
835	17	70.8	198	8	ADP30811	908	17	70.8	254	8	ADP30739	ADP30739	Human sec
836	17	70.8	198	8	ADP30477	909	17	70.8	254	8	ADP31396	ADP31396	Human sec
837	17	70.8	198	8	ADP30481	910	17	70.8	255	7	ADJ92145	ADJ92145	Human hai
838	17	70.8	200	7	ABO71047	911	17	70.8	255	8	ADP30658	ADP30658	Human sec
839	17	70.8	201	7	ABO60877	912	17	70.8	256	4	ABG08337	ABG08337	Human sec
840	17	70.8	203	7	ABO80007	913	17	70.8	256	6	ABU97138	ABU97138	Human
841	17	70.8	204	8	ADP31404	914	17	70.8	256	7	ABO75228	ABO75228	Pseudomon
842	17	70.8	206	7	ABO74022	915	17	70.8	256	8	ADN99494	ADN99494	Novel hum
843	17	70.8	208	4	AAU29506	916	17	70.8	257	5	ABP64573	ABP64573	Human ORF
844	17	70.8	208	5	ABG60794	917	17	70.8	258	8	ADP31084	ADP31084	Human sec
845	17	70.8	208	7	ABO80531	918	17	70.8	259	7	ABM8128	ABM8128	Rice abio
846	17	70.8	209	7	ABO77021	919	17	70.8	261	6	AAO27174	AAO27174	Pseudomon
847	17	70.8	210	8	ADP31563	920	17	70.8	263	3	AA60494	AA60494	Arabidops
848	17	70.8	210	8	ADP31562	921	17	70.8	263	7	ABO77554	ABO77554	Pseudomon
849	17	70.8	212	7	ABO67532	922	17	70.8	263	8	ABM81215	ABM81215	Tumour-as
850	17	70.8	215	7	ADP08255	923	17	70.8	264	8	ADP30788	ADP30788	Human sec
851	17	70.8	215	7	ABO72061	924	17	70.8	264	8	ADP31456	ADP31456	Human sec
852	17	70.8	215	9	ABM93776	925	17	70.8	264	8	ADP31412	ADP31412	Human sec
853	17	70.8	217	4	AAV97773	926	17	70.8	264	8	ADP31427	ADP31427	Human sec
854	17	70.8	218	7	ABO81158	927	17	70.8	266	8	ADP31427	ADP31427	Human sec
855	17	70.8	218	8	ADY23534	928	17	70.8	267	8	ADP327981	ADP327981	Novel hum
856	17	70.8	219	8	ADP31171	929	17	70.8	267	8	ADP31483	ADP31483	Human sec
857	17	70.8	219	8	ADY06385	930	17	70.8	270	8	ADP31217	ADP31217	Human sec
858	17	70.8	221	7	ADJ92147	931	17	70.8	270	8	ADP31321	ADP31321	Human sec
859	17	70.8	221	8	ADM87372	932	17	70.8	270	8	ADP31564	ADP31564	Human sec
860	17	70.8	222	5	ABM78572	933	17	70.8	270	9	ABE12365	ABE12365	Fusion pr
861	17	70.8	222	5	ABM78573	934	17	70.8	271	6	ABU6614	ABU6614	Protein e
862	17	70.8	223	3	ABM03862	935	17	70.8	271	6	ABU34590	ABU34590	Protein e
863	17	70.8	224	2	AAW22018	936	17	70.8	271	7	ADU92151	ADU92151	Human hai
864	17	70.8	224	2	ADW95939	937	17	70.8	272	6	ABU35898	ABU35898	Protein e
865	17	70.8	224	5	ADX15805	938	17	70.8	272	8	ADP12954	ADP12954	Protein e
866	17	70.8	225	5	ABM78570	939	17	70.8	273	8	ADP31236	ADP31236	Human sec
867	17	70.8	225	7	ABM88876	940	17	70.8	273	4	ABU94710	ABU94710	M. xanthu
868	17	70.8	225	8	ADP31501	941	17	70.8	274	4	AAU19418	AAU19418	Human dia
869	17	70.8	225	8	ADP30777	942	17	70.8	274	7	ABO72003	ABO72003	Pseudomon
870	17	70.8	227	7	ABO75458	943	17	70.8	275	8	ADP31679	ADP31679	Human sec
871	17	70.8	228	7	ADP30921	944	17	70.8	275	6	ABU34080	ABU34080	Protein e
872	17	70.8	231	5	ABP64726	945	17	70.8	276	7	ABO75093	ABO75093	Pseudomon
873	17	70.8	233	7	ABO75933	946	17	70.8	276	8	ADP31319	ADP31319	Human sec
874	17	70.8	235	7	ABO69554	947	17	70.8	276	8	ADP30584	ADP30584	Human sec
875	17	70.8	237	8	ADP31322	948	17	70.8	278	8	ADP31314	ADP31314	Human sec
876	17	70.8	238	9	ABE12367	949	17	70.8	279	8	ADP30846	ADP30846	Human sec
877	17	70.8	239	3	AAE21310	950	17	70.8	279	8	ADP31523	ADP31523	Human sec
878	17	70.8	239	8	ADY13038	951	17	70.8	279	8	ADP31503	ADP31503	Human sec
879	17	70.8	240	2	AAV22205	952	17	70.8	279	8	ADP31671	ADP31671	Human sec
880	17	70.8	240	8	ADP30737	953	17	70.8	279	8	ADP30605	ADP30605	Human sec
881	17	70.8	240	2	AAW44532	954	17	70.8	282	7	ADJ92143	ADJ92143	Human hai
882	17	70.8	244	2	AAW22985	955	17	70.8	282	7	ABO73521	ABO73521	Pseudomon
883	17	70.8	244	2	AAW51006	956	17	70.8	285	7	ABO73521	ABO73521	Pseudomon
884	17	70.8	244	3	ABE21323	957	17	70.8	288	8	ADP31463	ADP31463	Human sec
885	17	70.8	244	5	ABG63537	958	17	70.8	291	7	ABO68749	ABO68749	Pseudomon
886	17	70.8	244	6	AAE37572	959	17	70.8	293	6	ADAS5092	ADAS5092	Human pro
887	17	70.8	244	7	ADN80567	960	17	70.8	293	8	ADH09610	ADH09610	Human hos
888	17	70.8	244	7	ADN39212	961	17	70.8	294	8	ADP31076	ADP31076	Human sec
889	17	70.8	244	8	ADJ39734	962	17	70.8	294	8	ADP31473	ADP31473	Human sec
890	17	70.8	244	8	ADJ37158	963	17	70.8	297	8	ADP31192	ADP31192	Human sec
891	17	70.8	244	8	ADN04074	964	17	70.8	298	2	AAW43397	AAW43397	Human imi
892	17	70.8	244	8	ADN29289	965	17	70.8	302	3	AA607023	AA607023	Arabidops
893	17	70.8	244	8	ADN89076	966	17	70.8	304	3	ADP31659	ADP31659	Human sec
894	17	70.8	244	8	ADN72624	967	17	70.8	306	7	ADFO4021	ADFO4021	Bacterial
895	17	70.8	244	8	ADN72876	968	17	70.8	307	7	ADCO1287	ADCO1287	Enterohae
896	17	70.8	244	9	ADY67594	969	17	70.8	309	8	ADJ42146	ADJ42146	Plant tra
897	17	70.8	244	9	ADZ51357	970	17	70.8	310	4	ABBS5893	ABBS5893	Drosophi
898	17	70.8	245	7	ABO78304	971	17	70.8	311	4	ABB67853	ABB67853	Drosophi
899	17	70.8	246	8	ADP30619	972	17	70.8	311	4	ABG08405	ABG08405	Novel hum
900	17	70.8	246	9	ADY18495	973	17	70.8	312	1	AA650079	AA650079	T-cell an

```
974 17 70.8 312 1 AAP60471 Aap60471 Portion o
975 17 70.8 312 2 AAR53145 Aar53145 T-cell an
976 17 70.8 312 3 AAB25555 Aab25555 Pinus rad
977 17 70.8 312 4 AAG08404 Aag08404 Novel hum
978 17 70.8 312 8 ADP31505 Adp31505 Human sec
979 17 70.8 312 8 ADP30476 Adp30476 Human sec
980 17 70.8 313 7 ABM85966 Abm85966 Rice abio
981 17 70.8 315 5 AAE13348 Aae13348 Human TST
982 17 70.8 315 6 ABR58684 Abr58684 Human can
983 17 70.8 315 8 ADP31685 Adp31685 Human sec
984 17 70.8 316 3 AAY44862 Aay44862 Human tas
985 17 70.8 316 9 ADY18921 Ady18921 PRO polyP
986 17 70.8 317 8 ADP31615 Adp31615 Human sec
987 17 70.8 318 7 ABO78573 AbO78573 Pseudomon
988 17 70.8 318 8 ADP31066 Adp31066 Human sec
989 17 70.8 320 4 AAG31631 Aag31631 C glutami
990 17 70.8 325 3 AAB51703 Aab51703 Gene 34 h
991 17 70.8 326 3 AAG24798 Aag24798 Arabidops
992 17 70.8 336 8 ADP31320 Adp31320 Human sec
993 17 70.8 341 7 ABO77462 AbO77462 Pseudomon
994 17 70.8 343 7 ADF13953 Adf13953 Human end
995 17 70.8 345 3 AAG07022 Aag07022 Arabidops
996 17 70.8 345 3 AAG37538 Aag37538 Arabidops
997 17 70.8 345 8 ADP31204 Adp31204 Human sec
998 17 70.8 346 7 ABO81210 AbO81210 Pseudomon
999 17 70.8 346 7 ABM89803 Abm89803 Rice abio
1000 17 70.8 346 8 ADP31355 Adp31355 Human sec
```

## ALIGNMENTS

## RESULT 1

```
ADQ13022
ID ADQ13022 standard; peptide; 9 AA.
XX
AC ADQ13022;
XX
XX 07-OCT-2004 (first entry)
XX
DE Hepatitis B virus MHC class I restricted T-cell stimulating peptide 821.
XX
XX HBV; MHC class I restricted T-cell stimulating peptide; surface peptide;
XX core peptide; polymerase peptide; immunogenic composition; vaccine;
XX cytotoxic T-lymphocyte response; CTL response.
XX
OS Hepatitis B virus.
XX
PN WO2004058807-A2.
XX
PD 15-JUL-2004.
XX
PF 09-DEC-2003; 2003WO-EP013948.
XX
PR 24-DEC-2002; 2002EP-00447276.
XX
PA (ALGO-) ALGONOMICS NV.
XX
PI Laesters I, Desmet J, Stegmann T, Casteleijn B;
XX WPI; 2004-525861/50.
XX
DR WPI; 2004-525861/50.
XX
PT New peptide comprising a major histocompatibility complex class I
PT restricted T-cell stimulating epitope of the hepatitis B virus (HBV)
PT surface, core and/or polymerase polypeptide, for preparation of an HBV
PT immunogenic composition.
XX
XX Claim 5; SEQ ID NO 821; 108bp; English.
XX
XX The invention comprises major histocompatibility complex (MHC) class I
XX restricted T-cell stimulating epitopes of the hepatitis B virus (HBV)
XX surface, core and polymerase proteins. The peptides of the invention are
XX useful for generating an HBV immunogenic composition (e.g. vaccine) that
XX is useful for generating an HBV immunogenic composition (e.g. vaccine) that
```

CC induces a cytotoxic T-lymphocyte (CTL) response. The present amino acid  
CC sequence represents an MHC class I restricted T-cell stimulating HBV  
CC peptide of the invention.

SO Sequence 9 AA;

Query Match 75.0%; Score 18; DB 8; Length 9;  
Best Local Similarity 25.0%; Pred. No. 1.4e+06;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
Db 1 CASSSSSC 8

## RESULT 2

```
ADQ13021
ID ADQ13021 standard; peptide; 9 AA.
XX
AC ADQ13021;
XX
XX 07-OCT-2004 (first entry)
XX
DE Hepatitis B virus MHC class I restricted T-cell stimulating peptide 820.
XX
XX HBV; MHC class I restricted T-cell stimulating peptide; surface peptide;
XX core peptide; polymerase peptide; immunogenic composition; vaccine;
XX cytotoxic T-lymphocyte response; CTL response.
XX
OS Hepatitis B virus.
XX
PN WO2004058807-A2.
XX
PD 15-JUL-2004.
XX
PF 09-DEC-2003; 2003WO-EP013948.
XX
PR 24-DEC-2002; 2002EP-00447276.
XX
PA (ALGO-) ALGONOMICS NV.
XX
PI Laesters I, Desmet J, Stegmann T, Casteleijn B;
XX WPI; 2004-525861/50.
XX
DR WPI; 2004-525861/50.
XX
PT New peptide comprising a major histocompatibility complex class I
PT restricted T-cell stimulating epitope of the hepatitis B virus (HBV)
PT surface, core and/or polymerase polypeptide, for preparation of an HBV
PT immunogenic composition.
XX
XX Example 7; SEQ ID NO 820; 108bp; English.
XX
XX The invention comprises major histocompatibility complex (MHC) class I
XX restricted T-cell stimulating epitopes of the hepatitis B virus (HBV)
XX surface, core and polymerase proteins. The peptides of the invention are
XX useful for generating an HBV immunogenic composition (e.g. vaccine) that
XX induces a cytotoxic T-lymphocyte (CTL) response. The present amino acid
XX sequence represents an MHC class I restricted T-cell stimulating HBV
XX peptide of the invention.
XX
XX Sequence 9 AA;
XX
XX Query Match 75.0%; Score 18; DB 8; Length 9;
XX Best Local Similarity 25.0%; Pred. No. 1.4e+06;
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

RESULT 3  
AAW65943

ID AAW65943 standard; peptide; 11 AA.  
 XX AAW65943;  
 AC  
 XX 12-NOV-1998 (first entry)  
 XX  
 DE Molecule VI.  
 XX  
 KM Neurotrophin; cyclic; bicyclic; disulphide; nerve growth factor; NGF;  
 KM BDNF; NT-3; conformation; promoter.  
 XX  
 OS Synthetic.  
 OS Mammalia.  
 XX CA2205045-A.  
 XX  
 PD 12-MAY-1998.  
 XX  
 PF 09-MAY-1997; 97CA-02205045.  
 XX  
 PR 12-NOV-1996; 96CA-02190296.  
 XX  
 PA (TOOH ) UNIV QUEBENS KINGSTON.  
 PI Riopelle RJ, Dory MI, Shamovsky IL, Weaver DF, Ross GM;  
 XX WPI; 1998-457650/40.  
 DR  
 XX Inhibition of neurotrophin activity - using factor that interferes with  
 PT sub-unit interaction.  
 PS  
 XX Disclosure; Page 13; 103pp; English.  
 CC The invention relates to a method for reducing the biological activity of  
 CC a multimeric protein having at least 2 promoters. It comprises (a)  
 CC providing a factor that interacts with at least one portion of at least  
 CC one of the promoters which associates with a portion of the other  
 CC promoter in the absence of the factor; and (b) mixing the factor with the  
 CC multimeric protein so that the factor interacts with the portion(s) and  
 CC disrupts association of at least a portion of the promoters. The method  
 CC and compounds are useful for inhibiting neurotrophin-mediated activities  
 CC selected from neurotrophin receptor binding, neuron survival, neurite  
 CC outgrowth and epileptic effects. The present sequence is shown in the  
 CC specification  
 CC  
 SQ Sequence 11 AA;  
 QY 1 CXXXXXXC 8  
 Db 1 CAAAAAAC 8  
 75.0%; Score 18; DB 2; Length 11;  
 Best Local Similarity 25.0%; Pred. No. 2.3e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

RESULT 4  
 ID AAW65944 standard; peptide; 11 AA.  
 XX AAW65944;  
 AC  
 XX 12-NOV-1998 (first entry)  
 XX  
 DE Molecule XV.  
 XX  
 KM Neurotrophin; cyclic; bicyclic; disulphide; nerve growth factor; NGF;  
 KM BDNF; NT-3; conformation; promoter.  
 XX  
 OS Synthetic.  
 OS Mammalia.  
 XX CA2205045-A.  
 PN

XX 12-MAY-1998.  
 PD  
 XX  
 PF 09-MAY-1997; 97CA-02205045.  
 XX  
 PR 12-NOV-1996; 96CA-02190296.  
 XX  
 PA (TOOH ) UNIV QUEBENS KINGSTON.  
 PI Riopelle RJ, Dory MI, Shamovsky IL, Weaver DF, Ross GM;  
 XX WPI; 1998-457650/40.  
 DR  
 XX Inhibition of neurotrophin activity - using factor that interferes with  
 PT sub-unit interaction.  
 PS  
 XX Disclosure; Page 13; 103pp; English.  
 CC The invention relates to a method for reducing the biological activity of  
 CC a multimeric protein having at least 2 promoters. It comprises (a)  
 CC providing a factor that interacts with at least one portion of at least  
 CC one of the promoters which associates with a portion of the other  
 CC promoter in the absence of the factor; and (b) mixing the factor with the  
 CC multimeric protein so that the factor interacts with the portion(s) and  
 CC disrupts association of at least a portion of the promoters. The method  
 CC and compounds are useful for inhibiting neurotrophin-mediated activities  
 CC selected from neurotrophin receptor binding, neuron survival, neurite  
 CC outgrowth and epileptic effects. The present sequence is shown in the  
 CC specification  
 CC  
 SQ Sequence 11 AA;  
 QY 1 CXXXXXXC 8  
 Db 4 CAAAAAAC 11  
 75.0%; Score 18; DB 2; Length 11;  
 Best Local Similarity 25.0%; Pred. No. 2.3e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

RESULT 5  
 ID ADV23688 standard; peptide; 15 AA.  
 XX ADV23688;  
 AC  
 XX 10-MAR-2005 (first entry)  
 DT  
 DE HBV immunogenic peptide #489.  
 XX  
 KM Vaccine; virucide; antigen; autoimmune disease; infection;  
 KM Immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;  
 KM breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;  
 KM pancreas tumor; stomach tumor; bladder tumor; kidney tumor;  
 KM hodgekin's lymphoma.  
 XX  
 OS Hepatitis B virus.  
 OS  
 PN WO2004108753-A1.  
 XX  
 PD 16-DEC-2004.  
 XX  
 PF 10-JUN-2004; 2004MO-AU000775.  
 XX  
 PR 10-JUN-2003; 2003AU-00902875.  
 PR 25-MAR-2004; 2004AU-00901589.  
 XX  
 PA (UYME ) UNIV MELBOURNE.  
 PI Kent SJ;  
 XX  
 DR WPI; 2005-031657/03.

XX Use of at least one set of peptides in the preparation of a medicament  
PT for modulating an immune response, and for treating cancer or yeast,  
PT viral, bacterial, protozoal and mycoplasma infections.

XX Disclosure; SEQ ID NO 2108; 645pp; English.

XX The invention relates to the use of at least one set of peptides in the  
CC preparation of a medicament for modulating an immune response, where  
CC individual peptides of a respective set comprise different portions of an  
CC amino acid sequence corresponding to a single polypeptide of interest and  
CC display partial sequence identity or similarity to at least one other  
CC peptide of the same set of peptides (i.e., they are overlapping). Also  
CC included are an antigen-presenting cell which has been contacted with the  
CC peptides above and thus presents the peptides, a population of such  
CC antigen-presenting cells, a process for producing antigen-presenting  
CC cells for modulating an immune response to a polypeptide of interest, a  
CC method for producing antigen-specific lymphocytes, a composition  
CC comprising at least one set of the peptides (and a carrier and/or  
CC diluent), a method for modulating an immune response to a polypeptide of  
CC interest comprising administering to a patient in need at least one set  
CC of the peptides, a method for treatment and/or prophylaxis of a disease  
CC or condition associated with the presence of a polypeptide of interest  
CC and a composition of matter for modulating an immune response in a  
CC subject to a target antigen. The polypeptide of interest is also a  
CC disease- or condition-associated polypeptide that is a polypeptide  
CC produced by a pathogenic organism or a cancer, and produced by a  
CC pathogenic organism selected from yeast, viruses, bacteria, helminths,  
CC protozoans and mycoplasmas. The disease- or condition-associated  
CC polypeptide is produced by a cancer selected from melanoma, lung cancer,  
CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic  
CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant  
CC lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured  
CC antigen-presenting cells or their precursors are useful in the  
CC preparation of a medicament for the treatment of a disease or condition  
CC in a subject, which disease or condition is associated with the presence  
CC of aberrant expression of a target antigen, where the antigen-presenting  
CC cells or their precursors have not been subjected to activating  
CC conditions but have been contacted with an antigen that corresponds to  
CC the target antigen to express a processed or modified form of the antigen  
CC for presentation to the subject's immune system. The present sequence is  
CC one of a set of overlapping immunogenic peptides derived from a Hepatitis  
CC B virus protein.

XX Sequence 15 AA;

Query Match 75.0%; Score 18; DB 9; Length 15;  
Best Local Similarity 25.0%; Pred. No. 2.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
DB 4 CASSSSSC 11

RESULT 6  
ADV23687

ID ADV23687 standard; peptide; 15 AA.

XX ADV23687;

DT 10-MAR-2005 (first entry)

XX HBV immunogenic peptide #488.

XX Vaccine; virucide; antigen; autoimmune disease; infection;  
KM immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;  
KM breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;  
KM pancreas tumor; stomach tumor; bladder tumor; kidney tumor;  
KM hcgkin's lymphoma.

XX Hepatitis B virus.

XX

PN WO2004108753-A1.

XX 16-DEC-2004.

XX 10-JUN-2004; 2004WO-AU000775.

XX 10-JUN-2003; 2003AU-00902875.

XX 25-MAR-2004; 2004AU-00901589.

XX (UTME ) UNIV MELBOURNE.

XX Kent St;

XX WPI; 2005-031657/03.

PT Use of at least one set of peptides in the preparation of a medicament  
PT for modulating an immune response, and for treating cancer or yeast,  
PT viral, bacterial, protozoal and mycoplasma infections.

PS Disclosure; SEQ ID NO 2107; 645pp; English.

XX The invention relates to the use of at least one set of peptides in the  
CC preparation of a medicament for modulating an immune response, where  
CC individual peptides of a respective set comprise different portions of an  
CC amino acid sequence corresponding to a single polypeptide of interest and  
CC display partial sequence identity or similarity to at least one other  
CC peptide of the same set of peptides (i.e., they are overlapping). Also  
CC included are an antigen-presenting cell which has been contacted with the  
CC peptides above and thus presents the peptides, a population of such  
CC antigen-presenting cells, a process for producing antigen-presenting  
CC cells for modulating an immune response to a polypeptide of interest, a  
CC method for producing antigen-specific lymphocytes, a composition  
CC comprising at least one set of the peptides (and a carrier and/or  
CC diluent), a method for modulating an immune response to a polypeptide of  
CC interest comprising administering to a patient in need at least one set  
CC of the peptides, a method for treatment and/or prophylaxis of a disease  
CC or condition associated with the presence of a polypeptide of interest  
CC and a composition of matter for modulating an immune response in a  
CC subject to a target antigen. The polypeptide of interest is also a  
CC disease- or condition-associated polypeptide that is a polypeptide  
CC produced by a pathogenic organism or a cancer, and produced by a  
CC pathogenic organism selected from yeast, viruses, bacteria, helminths,  
CC protozoans and mycoplasmas. The disease- or condition-associated  
CC polypeptide is produced by a cancer selected from melanoma, lung cancer,  
CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic  
CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant  
CC lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured  
CC antigen-presenting cells or their precursors are useful in the  
CC preparation of a medicament for the treatment of a disease or condition  
CC in a subject, which disease or condition is associated with the presence  
CC of aberrant expression of a target antigen, where the antigen-presenting  
CC cells or their precursors have not been subjected to activating  
CC conditions but have been contacted with an antigen that corresponds to  
CC the target antigen to express a processed or modified form of the antigen  
CC for presentation to the subject's immune system. The present sequence is  
CC one of a set of overlapping immunogenic peptides derived from a Hepatitis  
CC B virus protein.

XX Sequence 15 AA;

Query Match 75.0%; Score 18; DB 9; Length 15;  
Best Local Similarity 25.0%; Pred. No. 2.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
DB 8 CASSSSSC 15

RESULT 7  
ABB42273

ID ABB42273 standard; peptide; 24 AA.

XX

```
AC ABB42273;
XX
DT 04-FEB-2002 (first entry)
DE Peptide #9779 encoded by human foetal liver single exon probe.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
PN WO200157277-A2.
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn Sq, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 34908; 639bp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 24 AA;
XX
Query Match 75.0%; Score 18; DB 4; Length 24;
Best Local Similarity 25.0%; Pred. No. 2.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0.
OY 1 CXXXXXXC 8
Db |
|
12 CTSTTTC 19
XX
RESULT 8
ID AAM36081
AA AAM36081 standard; protein; 24 AA.
XX
AC AAM36081;
XX
DT 17-OCT-2001 (first entry)
DE Peptide #10118 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
OS Homo sapiens.
XX
PW WO200157272-A2.
XX
PD 09-AUG-2001.
```

[illegible]

DR WPI; 2001-488900/53.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human bone marrow.  
XX  
PS Example 4; SEQ ID NO 36279; 658bp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention  
SQ Sequence 24 AA;  
  
Query Match 75.0%; Score 18; DB 4; Length 24;  
Best Local Similarity 25.0%; Pred. No. 2.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
OY 1 CXXXXXXC 8  
12 CTTSTTTC 19  
DB  
  
RESULT 10  
AAM63159  
ID AAM63159 standard; protein; 24 AA.  
XX  
AC AAM63159;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35264.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe; microarray;  
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.  
XX  
OS Homo sapiens.  
XX  
PN W0200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000667.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains.  
XX  
PS Example 4; SEQ ID NO 35264; 650bp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention  
XX

SQ Sequence 24 AA;  
  
Query Match 75.0%; Score 18; DB 4; Length 24;  
Best Local Similarity 25.0%; Pred. No. 2.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
OY 1 CXXXXXXC 8  
12 CTTSTTTC 19  
DB  
  
RESULT 11  
ABG57702  
ID ABG57702 standard; peptide; 24 AA.  
XX  
AC ABG57702;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human liver peptide, SEQ ID No 36350.  
XX  
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
KW hypercholesterolaemia; coronary heart disease.  
XX  
OS Homo sapiens.  
XX  
PN W0200157273-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000664.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488898/53.  
XX  
PF Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human adult liver.  
XX  
PS Claim 27; SEQ ID NO 36350; 658bp; English.  
XX  
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult liver.  
CC (I) may be used for predicting, measuring and displaying gene expression  
CC in samples derived from human adult liver. The genes identified may be  
CC involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
CC associated with coronary heart disease. ABG47348-ABG59350 represent human  
CC liver single exon encoded peptides of the invention. Note: The sequence  
CC information for this patent does not appear in the printed specification  
CC but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 24 AA;  
  
Query Match 75.0%; Score 18; DB 4; Length 24;  
Best Local Similarity 25.0%; Pred. No. 2.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
OY 1 CXXXXXXC 8

Db 12 CTTSTTTC 19

## RESULT 12

ID ABB68695 standard; protein; 69 AA.

AC ABB68695;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 32877.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-6556860/75.

XX N-PSDB; ABL12798.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

XX Disclosure; SEQ ID NO 32877; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (AB57737-AB572072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pcc\_sequences

XX SQ Sequence 69 AA;

Query Match 75.0%; Score 18; DB 4; Length 69;

Best Local Similarity 25.0%; Pred. No. 3.3e+02; Mismatches 0; Gaps 0;

Matches 2; Conservative 0; Indels 0;

QY 1 CXXXXXXC 8

Db 9 CSSAASC 16

## RESULT 13

ID ADP30715 standard; protein; 69 AA.

AC ADP30715;

DT 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #1482.

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

XX WO2004035732-A2.

XX 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

XX 17-SEP-2002; 2002US-0411048P.

XX 17-SEP-2002; 2002US-0411052P.

XX 17-SEP-2002; 2002US-0411055P.

XX 17-SEP-2002; 2002US-0411073P.

XX 17-SEP-2002; 2002US-0411082P.

XX 17-SEP-2002; 2002US-0411101P.

XX 17-SEP-2002; 2002US-0411111P.

XX 18-APR-2003; 2003US-0463700P.

XX 18-APR-2003; 2003US-0463708P.

XX 18-APR-2003; 2003US-0463716P.

XX 18-APR-2003; 2003US-0463732P.

XX 02-MAY-2003; 2003US-0467189P.

XX 02-MAY-2003; 2003US-0467201P.

XX 02-MAY-2003; 2003US-0467203P.

XX 02-MAY-2003; 2003US-0467230P.

XX 19-MAY-2003; 2003US-0471306P.

XX 19-MAY-2003; 2003US-0471336P.

XX 22-MAY-2003; 2003US-0472420P.

XX 22-MAY-2003; 2003US-0472430P.

XX 09-JUN-2003; 2003US-0476609P.

XX 09-JUN-2003; 2003US-0476641P.

XX 08-JUL-2003; 2003US-0485218P.

XX 08-JUL-2003; 2003US-0485223P.

XX 08-JUL-2003; 2003US-0485224P.

XX 08-JUL-2003; 2003US-0485325P.

XX 14-JUL-2003; 2003US-0486446P.

XX 14-JUL-2003; 2003US-0486480P.

DR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
PS  
PS Claim 1; SEQ ID NO 2713; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.  
XX  
SQ Sequence 69 AA;

Query Match 75.0%; Score 18; DB 8; Length 69;  
Best Local Similarity 25.0%; Pred. No. 3.3e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 18 CTTTAAC 25

RESULT 14  
ABB39058  
ID ABB39058 standard; peptide; 91 AA.  
XX  
AC ABB39058;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #6564 encoded by human foetal liver single exon probe.  
XX  
KM Human, foetal liver; gene expression; single exon nucleic acid probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000669.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human foetal liver.  
XX  
PS  
PS Claim 27; SEQ ID NO 31693; 639bp + Sequence listing; English.  
XX

CC The invention relates to a single exon nucleic acid probe for measuring  
CC human gene expression in a sample derived from human foetal liver. The  
CC single exon nucleic acid probes may be used for predicting, measuring and  
CC displaying gene expression in samples derived from human foetal liver. The  
CC present sequence is a peptide encoded by a single exon nucleic acid probe  
CC of the invention. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX

SQ Sequence 91 AA;  
Query Match 75.0%; Score 18; DB 4; Length 91;  
Best Local Similarity 25.0%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 29 CSSSSSSC 36

RESULT 15  
AAM32549  
ID AAM32549 standard; protein; 91 AA.  
XX  
AC AAM32549;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Peptide #6586 encoded by probe for measuring placental gene expression.

XX  
KM Probe; microarray; human; placenta; antenatal diagnosis;  
XX genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX

PF 30-JAN-2001; 2001WO-US000663.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488897/53.  
XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human placenta.  
XX

PS Claim 27; SEQ ID NO 32818; 654bp; English.

XX  
CC The present invention relates to single exon nucleic acid probes (SENP:  
CC see A113315-A157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders



```
XX
SQ Sequence 91 AA;
Query Match 75.0%; Score 18; DB 4; Length 91;
Best Local Similarity 25.0%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
Db 29 CSSSSSSC 36
RESULT 16
AAM72290
ID AAM72290 standard; protein; 91 AA.
XX
AC AAM72290;
XX
AC AAM72290;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 32596.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 32596; 658bp + Sequence listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
SQ Sequence 91 AA;
Query Match 75.0%; Score 18; DB 4; Length 91;
Best Local Similarity 25.0%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
Db 29 CSSSSSSC 36
RESULT 17
AAM59710
ID AAM59710 standard; protein; 91 AA.
```

```
XX
AC AAM59710;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 31815.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brain.
XX
PS Example 4; SEQ ID NO 31815; 650bp + Sequence listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
SQ Sequence 91 AA;
Query Match 75.0%; Score 18; DB 4; Length 91;
Best Local Similarity 25.0%; Pred. No. 3.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CXXXXXXC 8
Db 29 CSSSSSSC 36
RESULT 18
ABG53976
ID ABG53976 standard; peptide; 91 AA.
XX
AC ABG53976;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver peptide; SEQ ID NO 32624.
XX
KW Human; liver; cirrhosis; hyperlipoproteinemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
```

PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000664.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-488898/53.  
 XX  
 DR  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human adult liver.  
 XX  
 PS Claim 27; SEQ ID NO 32624; 658bp; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult liver.  
 CC (I) may be used for predicting, measuring and displaying gene expression  
 CC in samples derived from human adult liver. The genes identified may be  
 CC involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human  
 CC liver single exon encoded peptides of the invention. Note: The sequence  
 CC information for this patent does not appear in the printed specification  
 CC but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 XX Sequence 91 AA;  
 SQ  
 Query Match 75.0%; Score 18; DB 4; Length 91;  
 Best Local Similarity 25.0%; Pred. No. 3.5e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Oy 1 CXXXXXXC 8  
 Db 29 CSSSSSSC 36  
 RESULT 19  
 ABG42105  
 ID ABG42105 standard; peptide: 91 AA.  
 XX  
 AC ABG42105;  
 XX  
 DT 19-AUG-2002 (first entry)  
 XX  
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 31770.  
 XX  
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberculous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 XX  
 XX Homo sapiens.  
 XX  
 XX WO200186003-A2.  
 PN  
 XX

PD 15-NOV-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000665.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2002-114183/15.  
 XX  
 DR  
 PT Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples.  
 XX  
 PS Claim 27; SEQ ID NO 31770; 634bp; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of probes  
 CC; the novel set of probes which hybridise at high stringency to a nucleic  
 CC acid expressed in the human lung; measuring gene expression in a sample  
 CC collected from human lung, comprising (a) contacting the array with a  
 CC collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of the  
 CC array; identifying exons in a eukaryotic genome, comprising (a)  
 CC algorithmically predicting at least one exon from genomic sequences of  
 CC the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene expression  
 CC analysis, and for identifying exons in a gene, particularly using human  
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberculous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a peptide/protein encoded by a single exon probe of  
 CC the invention. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 XX Sequence 91 AA;  
 SQ  
 Query Match 75.0%; Score 18; DB 5; Length 91;  
 Best Local Similarity 25.0%; Pred. No. 3.5e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Oy 1 CXXXXXXC 8  
 Db 29 CSSSSSSC 36

RESULT 20  
ADP30859  
ID ADP30859 standard; protein; 93 AA.  
XX AC ADP30859;  
XX DT 12-AUG-2004 (first entry)  
XX DE Human secreted protein SEQ ID #1626.  
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX KM cancer; inflammatory; immune; human secreted protein.  
XX OS Homo sapiens.  
XX PN WO2004035732-A2.  
XX PD 29-APR-2004.  
XX PF 28-AUG-2003; 2003WO-US026780.  
XX PR 29-AUG-2002; 2002US-0406576P.  
XX PR 29-AUG-2002; 2002US-0406579P.  
XX PR 29-AUG-2002; 2002US-0406585P.  
XX PR 29-AUG-2002; 2002US-0406588P.  
XX PR 29-AUG-2002; 2002US-0406608P.  
XX PR 29-AUG-2002; 2002US-0406611P.  
XX PR 29-AUG-2002; 2002US-0406612P.  
XX PR 29-AUG-2002; 2002US-0406616P.  
XX PR 29-AUG-2002; 2002US-0406640P.  
XX PR 29-AUG-2002; 2002US-0406642P.  
XX PR 29-AUG-2002; 2002US-0406653P.  
XX PR 29-AUG-2002; 2002US-0406655P.  
XX PR 29-AUG-2002; 2002US-0406666P.  
XX PR 17-SEP-2002; 2002US-0410946P.  
XX PR 17-SEP-2002; 2002US-0410947P.  
XX PR 17-SEP-2002; 2002US-0410948P.  
XX PR 17-SEP-2002; 2002US-0410949P.  
XX PR 17-SEP-2002; 2002US-0410953P.  
XX PR 17-SEP-2002; 2002US-0410957P.  
XX PR 17-SEP-2002; 2002US-0410958P.  
XX PR 17-SEP-2002; 2002US-0410959P.  
XX PR 17-SEP-2002; 2002US-0410960P.  
XX PR 17-SEP-2002; 2002US-0410961P.  
XX PR 17-SEP-2002; 2002US-0410962P.  
XX PR 17-SEP-2002; 2002US-0411019P.  
XX PR 17-SEP-2002; 2002US-0411023P.  
XX PR 17-SEP-2002; 2002US-0411024P.  
XX PR 17-SEP-2002; 2002US-0411032P.  
XX PR 17-SEP-2002; 2002US-0411035P.  
XX PR 17-SEP-2002; 2002US-0411037P.  
XX PR 17-SEP-2002; 2002US-0411041P.  
XX PR 17-SEP-2002; 2002US-0411045P.  
XX PR 17-SEP-2002; 2002US-0411046P.  
XX PR 17-SEP-2002; 2002US-0411048P.  
XX PR 17-SEP-2002; 2002US-0411052P.  
XX PR 17-SEP-2002; 2002US-0411055P.  
XX PR 17-SEP-2002; 2002US-0411073P.  
XX PR 17-SEP-2002; 2002US-0411082P.  
XX PR 17-SEP-2002; 2002US-0411101P.  
XX PR 18-APR-2003; 2003US-0463700P.  
XX PR 18-APR-2003; 2003US-0463708P.  
XX PR 18-APR-2003; 2003US-0463716P.  
XX PR 18-APR-2003; 2003US-0463732P.  
XX PR 02-MAY-2003; 2003US-0467199P.  
XX PR 02-MAY-2003; 2003US-0467201P.  
XX PR 02-MAY-2003; 2003US-0467203P.  
XX PR 19-MAY-2003; 2003US-0467230P.  
XX PR 19-MAY-2003; 2003US-0471306P.  
XX PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halebeck RF, Huang MM, Kothakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX DR WPI; 2004-348438/32.  
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX PT genetic, bacterial and viral diseases.  
XX PS Claim 1; SEQ ID NO 2857; 428bp; English.  
XX CC The present invention relates to an isolated nucleic acid molecule  
XX CC encoding a polypeptide which is believed to be cyostatic,  
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX CC composition and methods are useful for diagnosing, preventing and  
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present  
XX CC sequence represents a human secreted protein. The present sequence is  
XX CC available on WIPOMER and is not in the specification.  
SQ Sequence 93 AA;  
Query Match 75.0%; Score 18; DB 8; Length 93;  
Best Local Similarity 25.0%; Pred.No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
CY 1 CXXXXXXC 8  
DB 85 CATATATC 92  
RESULT 21  
ABO83167  
ID ABO83167 standard; protein; 99 AA.  
XX AC ABO83167;  
XX DT 29-JUL-2004 (first entry)  
XX DE Pseudomonas aeruginosa polypeptide #15342.  
XX DE Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX KW Pseudomonas aeruginosa.  
XX OS Pseudomonas aeruginosa.  
XX PN US6551795-B1.  
XX PD 22-APR-2003.  
XX PF 18-FEB-1999; 99US-00252991.  
XX PR 18-FEB-1998; 98US-0074788P.  
XX PR 27-JUL-1998; 98US-0094190P.



CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytosolic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer, inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.  
XX  
SQ Sequence 99 AA;

Query Match 75.0%; Score 18; DB 8; Length 99;  
Best Local Similarity 25.0%; Pred. No. 3.6e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 77 CAAATATC 84

## RESULT 23

ID ABO74878 standard; protein; 107 AA.

XX ABO74878;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #7053.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI: 2003-615309/58.

XX N-PSDB; ABD08449.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 23624; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX

SQ Sequence 107 AA;

Query Match 75.0%; Score 18; DB 7; Length 107;  
Best Local Similarity 25.0%; Pred. No. 3.6e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 29 CSASSSAC 36

## RESULT 24

ID ABO73394 standard; protein; 115 AA.

XX ABO73394;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #5569.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

DR WPI: 2003-615309/58.

XX N-PSDB; ABD06965.

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 22140; 455pp; English.

CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX

SQ Sequence 115 AA;

Query Match 75.0%; Score 18; DB 7; Length 115;  
Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 83 CSAATSTC 90

RESULT 25  
ADP04777  
ID ADF04777 standard; protein, 122 AA.  
XX  
AC ADF04777;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE Bacterial polypeptide #890.  
XX  
KW Proteus mirabilis infection; bacterial infection; antibacterial;  
KM immunostimulant.  
XX  
OS Proteus mirabilis.  
XX  
PN US6605709-B1.  
PD 12-AUG-2003.  
XX  
PF 05-APR-2000; 2000US-00543681.  
XX  
PR 09-APR-1999; 99US-0128706P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Breton GL;  
XX  
DR WPI; 2003-895291/82.  
XX  
N-PSDB; ADF00605.  
PT New Proteus mirabilis polypeptides and polynucleotides, useful as  
PT reagents for diagnosis of bacterial disease, as components of  
PT antibacterial vaccines, as targets for antibacterial drugs, or as  
PT biocontrol agents for plants.  
XX  
XX  
XX  
PS Disclosure; SEQ ID NO 5062; 870pp; English.  
XX  
XX  
CC The invention relates to new Proteus mirabilis polypeptides and  
CC polynucleotides. The invention also relates to antibodies against the  
CC polypeptides, methods for producing the polypeptides, a method of  
CC generating vaccines for immunising an individual against P. mirabilis, a  
CC method for evaluating a compound for the ability to bind a P. mirabilis  
CC polypeptide and a method for screening test compounds for anti-bacterial  
CC activity. The polypeptides and polynucleotides are useful as molecular  
CC targets for diagnosing, preventing and treating pathological conditions  
CC resulting from bacterial infection, as reagents for diagnosis of  
CC bacterial diseases, as components of antibacterial vaccines, as targets  
CC for antibacterial drugs or as bio-control agents for plants. This  
CC sequence represents a Proteus mirabilis polypeptide of the invention.  
XX  
SQ Sequence 122 AA;  
Query Match 75.0%; Score 18; DB 7; Length 122;  
Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476611P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493572P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PA  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Haisan L, Lannemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
DR  
XX  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
XX Claim 1; SEQ ID NO 2701; 428bp; English.  
PS  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMER and is not in the specification.  
XX  
XX Sequence 135 AA;  
SQ

Query Match 75.0%; Score 18; DB 8; Length 135;  
Best Local Similarity 25.0%; Pred. No. 3.8e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
DB 105 CTTTATC 112

RESULT 27  
ABO75730  
ID ABO75730 standard; protein; 144 AA.  
XX  
XX ABO75730;  
AC  
XX  
XX 29-JUL-2004 (first entry)  
DT  
XX  
XX Pseudomonas aeruginosa polypeptide #7905.  
DB  
XX  
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
KW  
XX  
XX Pseudomonas aeruginosa.  
OS  
XX  
XX US651795-B1.  
PN  
XX  
XX 22-APR-2003.  
PD  
XX  
XX 18-FEB-1999; 99US-00252991.  
PF  
XX  
XX 18-FEB-1998; 98US-0074788P.  
PR  
XX  
XX 27-JUL-1998; 98US-0094190P.  
PR  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
PA  
XX  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
PI  
XX  
XX WPI; 2003-615309/58.  
DR  
XX  
XX N-PSDB; ABD09301.  
DR  
XX  
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of

PT pathological conditions resulting from bacterial infection.  
XX  
XX  
XX Disclosure; SEQ ID NO 24476; 455bp; English.  
PS  
XX  
XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using bioclip technology. Sequences ABO67828-  
CC ABO64336 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
XX Sequence 144 AA;  
SQ

Query Match 75.0%; Score 18; DB 7; Length 144;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
DB 36 CSATSAC 43

RESULT 28  
ADP31453  
ID ADP31453 standard; protein; 144 AA.  
XX  
XX ADP31453;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
DT  
XX  
XX Human secreted protein SEQ ID #2220.  
DE  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2004035732-A2.  
PN  
XX  
XX 29-APR-2004.  
PD  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406579P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406585P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406588P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406608P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406611P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406612P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406616P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406642P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406646P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406653P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406655P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406666P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410946P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410947P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410948P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410949P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410953P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-048646P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Heestir K, Beaurang PA, Behrens D;  
PI Haldenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3451; 428pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present  
XX sequence represents a human secreted protein. The present sequence is  
XX available on WIPOWEB and is not in the specification.  
XX  
XX Sequence 144 AA;

Query Match 75.0%; Score 18; DB 8; Length 144;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 CXXXXXXC 8  
Db 20 CTATATAC 27  
RESULT 29  
ADX90450  
ID ADX90450 standard; protein; 144 AA.  
XX  
AC ADX90450;  
XX  
DT 21-APR-2005 (first entry)  
XX  
DE Plant full length insert polypeptide seqid 53114.  
XX  
KW plant protectant; plant growth regulant; gene therapy; plant;  
KW recombinant DNA construct; physical array; plant breeding marker;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
KW extreme osmotic condition; pathogen tolerance; pest tolerance;  
KW growth rate; cell cycle pathway; disease resistance;  
KW galactomannan production; lignin production; plant growth regulator;  
KW yield; plant growth; plant development; seed oil; protein yield;  
KW protein content.  
XX  
OS unidentified.  
XX  
XX US2004034888-A1.  
XX  
XX 19-FEB-2004.  
XX  
XX 28-APR-2003; 2003US-00425114.  
XX  
XX 06-MAY-1999; 99US-00304517.  
XX  
XX 05-NOV-2001; 2001US-00985678.  
XX  
XX (LITU/) LITU J.  
XX (ZHOU/) ZHOU Y.  
XX (KOVA/) KOVALIC D K.  
XX (SCRE/) SCREEN S E.  
XX (TABAS/) TABASKA J E.  
XX (CAOV/) CAO Y.  
XX  
XX Liu J, Zhou Y, Kovajic DK, Screen SF, Tabaska JE, Cao Y;  
XX  
XX WPI; 2004-180133/17.  
XX  
XX New recombinant DNA construct, useful for improving plant tolerance to  
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
XX pests, for conferring increased resistance to plant disease, or for  
XX improving yield.  
XX  
XX Claim 1; SEQ ID NO 53114; 15pp; English.  
XX  
XX The invention describes a recombinant DNA construct comprising a  
XX polynucleotide consisting of a sequence encoding an amino acid sequence  
XX available in electronic form from the US patent office at  
XX ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide  
XX of the invention are also useful in physical arrays of molecules and as  
XX plant breeding markers. The recombinant DNA construct is useful for  
XX improving plant tolerance to cold, heat, drought, herbicides, extreme  
XX osmotic conditions, pathogens or pests, for manipulating growth rate in  
XX plant cells by modification of the cell cycle pathway, for conferring  
XX increased resistance to plant disease, for producing galactomannan,  
XX lignin or plant growth regulators, for increasing the rate of homologous  
XX recombination in plants, for improving yield by modification of  
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake  
XX or by providing improved plant growth and development under at least one  
XX stress condition or for modifying seed oil or protein yield and/or



CC content. This is the amino acid sequence of a plant full length insert  
CC polypeptide that can be used in the recombinant DNA construct of the  
CC invention.  
XX  
SQ Sequence 144 AA;  
Query Match 75.0%; Score 18; DB 8; Length 144;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 CXXXXXXC 8  
DB 4 CTSSSAC 11  
RESULT 30  
ABU45351  
ID ABU45351 standard; protein; 150 AA.  
AC ABU45351;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by prokaryotic essential gene #30878.  
XX  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS Salmoneella paratyphi.  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
XX  
PR 06-SEP-2001; 2001US-00948993.  
XX  
PR 25-OCT-2001; 2001US-0342923P.  
XX  
PR 08-FEB-2002; 2002US-00072851.  
XX  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
DR WPI; 2003-029926/02.  
XX  
DR N-PSDB; ACA49221.  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 73275; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 623 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 150 AA;  
Query Match 75.0%; Score 18; DB 6; Length 150;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 CXXXXXXC 8  
DB 16 CATTAASC 23  
RESULT 31  
AAB41043  
ID AAB41043 standard; protein; 153 AA.  
AC AAB41043;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Human ORF807 polypeptide sequence SEQ ID NO:1614.  
XX  
XX Human; open reading frame; ORF; detection; cytostatic; hepatotropic;  
KW vulnereary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW immunostimulant; thrombolytic; coagulable; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
KW antiviral; antibacterial; antifungal; antineumatic; antihydroid;  
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
KW thrombosis; contraceptive.  
XX  
XX Homo sapiens.  
XX  
XX WO200058473-A2.  
XX  
XX PD 05-OCT-2000.  
XX  
XX PF 31-MAR-2000; 2000WO-US008621.  
XX  
XX PR 31-MAR-1999; 99US-0127607P.  
XX  
XX PR 02-APR-1999; 99US-0127636P.  
XX  
XX PR 05-APR-1999; 99US-0127728P.  
XX  
XX PR 30-MAR-2000; 2000US-00540763.  
XX  
XX (CIRRA-) CIRAGEN CORP.  
XX  
XX Shimkets RA, Leach M;  
XX  
XX WPI; 2000-602362/57.  
XX  
XX DR N-PSDB; AAC75252.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
XX useful for treating e.g. cancers, proliferative disorders,  
XX neurodegenerative disorders and cardiovascular disease.  
XX  
XX PT

XX Claim 11; Page 1302; 5507pp; English.  
PS  
XX  
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytostatic; hepatotropic; valineary;  
CC antiproliferative; antiparkinsonian; neurotrophic; neuroprotective; osteopathic;  
CC anticovulsant; antichrithic; immunosuppressant; immunostimulant;  
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;  
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;  
CC antiviral; antifungal; antirheumatic; antichryoid; and antinaemic. The  
CC sequences can be used for determining the presence of or predisposition  
CC to, or preventing or treating pathological conditions associated with an  
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX  
CC proteins in gene therapy vectors. The proteins and nucleic acids may be  
CC used to treat cancers, proliferative disorders, neurodegenerative  
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
CC storage, systemic lupus erythematosus, severe combined immunodeficiency  
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to  
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive  
XX  
SQ Sequence 153 AA;  
Query Match 75.0%; Score 18; DB 3; Length 153;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 CXXXXXXC 8  
Db 78 CSSSASC 85  
RESULT 32  
ABP09251  
ID ABP09251 standard; protein; 153 AA.  
XX  
AC ABP09251;  
DT 25-JUN-2002 (first entry)  
XX  
DE Human ORFX protein sequence SEQ ID NO:18484.  
XX  
KW Human, open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
KW hypertension; hypothyroidism; cholesterol ester storage disease;  
KW immune deficiency; immune disorder; infectious disease;  
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
KW myaesthesia gravis.  
XX  
OS Homo sapiens.  
XX  
PN MO200192523-A2.  
XX  
PD 06-DEC-2001.  
XX  
PF 29-MAY-2001; 2001WO-US010836.  
XX  
PR 30-MAY-2000; 2000US-0206132P.  
PR 29-AUG-2000; 2000US-0228716P.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Shinkets RA, Leach MD;  
XX  
DR WPI; 2002-106308/14.  
XX  
PT N-PSDB; ABN25003.  
Novel human polypeptides and polynucleotides useful for diagnosing,

PT preventing and treating cardiovascular disease, neurodegenerative,  
PT hyperproliferative disorders and autoimmune disorders.  
PS  
XX  
XX Disclosure; SEQ ID NO 18484; 1037pp; English.  
XX  
CC The present invention describes substantially purified human proteins  
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
CC in the specification)). ABN15762 to ABN27252 encode the human ORFX  
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
CC treating or preventing a pathology associated with an ORFX-associated  
CC disorder in humans, and in the manufacture of a medicament for treating a  
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
CC sequences can be used in gene therapy. ORFX sequences can be used in the  
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
CC storage disease, various immune deficiencies and disorders, infectious  
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
CC arthritis, autoimmune thyroiditis, myaesthesia gravis, graft-versus-host  
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
CC bone degenerative disorders, or periodontal disease, and for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues and conditions resulting from  
CC systemic cytokine damage. N.B. The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 153 AA;  
Query Match 75.0%; Score 18; DB 5; Length 153;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 CXXXXXXC 8  
Db 78 CSSSASC 85  
RESULT 33  
ABO74421  
ID ABO74421 standard; protein; 154 AA.  
XX  
AC ABO74421;  
DT 29-JUL-2004 (first entry)  
XX  
DE Pseudomonas aeruginosa polypeptide #6596.  
XX  
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX  
OS Pseudomonas aeruginosa.  
XX  
PN US6551795-B1.  
XX  
PD 22-APR-2003.  
XX  
PF 18-FEB-1999; 99US-00252991.  
XX  
PR 18-FEB-1998; 98US-0074788P.  
PR 27-JUL-1998; 98US-0094190P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX  
DR WPI; 2003-615309/58.  
XX  
PT N-PSDB; ABD07992.  
Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of

PT pathological conditions resulting from bacterial infection.  
XX  
XX Disclousure; SEQ ID NO 21167; 455pp; English.  
XX  
CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
XX Sequence 154 AA;  
SQ  
Query Match 75.0%; Score 18; DB 7; Length 154;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 CXXXXXXC 8  
Db 120 CAAAASAC 127  
RESULT 34  
ABO77836  
ID ABO77836 standard; protein; 156 AA.  
XX  
XX ABO77836;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
XX Pseudomonas aeruginosa polypeptide #10011.  
XX  
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX  
XX Pseudomonas aeruginosa.  
XX  
XX US6551795-B1.  
XX  
XX 22-APR-2003.  
XX  
XX 18-FEB-1999; 99US-00252991.  
XX  
XX 18-FEB-1998; 98US-0074788P.  
XX  
XX 27-JUL-1998; 98US-0094190P.  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX  
XX WPI; 2003-615309/58.  
XX  
XX N-PSDB; ABD11407.  
XX  
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX pathological conditions resulting from bacterial infection.  
XX  
XX Disclousure; SEQ ID NO 26582; 455pp; English.  
XX  
XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,

CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
XX Sequence 156 AA;  
SQ  
Query Match 75.0%; Score 18; DB 7; Length 156;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 CXXXXXXC 8  
Db 40 CSASSAC 47  
RESULT 35  
ADP30761  
ID ADP30761 standard; protein; 159 AA.  
XX  
XX ADP30761;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #1528.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; Immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX  
XX 29-AUG-2002; 2002US-0406579P.  
XX  
XX 29-AUG-2002; 2002US-0406585P.  
XX  
XX 29-AUG-2002; 2002US-0406588P.  
XX  
XX 29-AUG-2002; 2002US-0406608P.  
XX  
XX 29-AUG-2002; 2002US-0406611P.  
XX  
XX 29-AUG-2002; 2002US-0406616P.  
XX  
XX 29-AUG-2002; 2002US-0406640P.  
XX  
XX 29-AUG-2002; 2002US-0406642P.  
XX  
XX 29-AUG-2002; 2002US-0406646P.  
XX  
XX 29-AUG-2002; 2002US-0406653P.  
XX  
XX 29-AUG-2002; 2002US-0406655P.  
XX  
XX 29-AUG-2002; 2002US-0406666P.  
XX  
XX 17-SEP-2002; 2002US-0410946P.  
XX  
XX 17-SEP-2002; 2002US-0410947P.  
XX  
XX 17-SEP-2002; 2002US-0410948P.  
XX  
XX 17-SEP-2002; 2002US-0410949P.  
XX  
XX 17-SEP-2002; 2002US-0410953P.  
XX  
XX 17-SEP-2002; 2002US-0410957P.  
XX  
XX 17-SEP-2002; 2002US-0410958P.  
XX  
XX 17-SEP-2002; 2002US-0410959P.  
XX  
XX 17-SEP-2002; 2002US-0410960P.  
XX  
XX 17-SEP-2002; 2002US-0410962P.  
XX  
XX 17-SEP-2002; 2002US-0411019P.  
XX  
XX 17-SEP-2002; 2002US-0411022P.  
XX  
XX 17-SEP-2002; 2002US-0411023P.  
XX  
XX 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486961P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493577P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Heetir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang WM, Kothakota S, Haisan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2759; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cyostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
XX Sequence 159 AA;

Query Match 75.0%; Score 18; DB 8; Length 159;  
Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
DB 125 CAATATATC 132

RESULT 36  
ADP30758  
ID ADP30758 standard; protein; 162 AA.  
XX  
XX ADP30758;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
DT  
XX  
XX Human secreted protein SEQ ID #1525.  
DE  
XX  
XX Cyostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
OS  
XX WO2004035732-A2.  
FN  
XX  
XX 29-APR-2004.  
PD  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Heetir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Halsehan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2756; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytosolic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 162 AA;  
QY  
Db 1 CXXXXXXC 8  
53 CAATTATC 60  
Query Match 75.0%; Score 18; DB 8; Length 162;  
Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
RESULT 37  
ADP30756  
ID ADP30756 standard; protein; 162 AA.  
XX  
AC ADP30756;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1523.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX OS  
XX MO2004035732-A2.  
XX  
XX 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX

PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406610P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411025P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467189P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 02-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Heetir K, Beaurang PA, Behrens D;  
XX Halenbeck RF, Huang MM, Kothakota S, Halsehan L, Linnemann T;  
PI

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2754; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytosstatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 162 AA;  
  
Query Match 75.0%; Score 18; DB 8; Length 162;  
Best Local Similarity 25.0%; Pred. No. 4e+02; 6; Indels 0; Gaps 0;  
Matches 2; Conservative 0; Mismatches 0;  
  
QY 1 CXXXXXXC 8  
Db 53 CAATTATC 60  
  
RESULT 38  
ADP30757  
ID ADP30757 standard; protein; 162 AA.  
XX  
AC ADP30757;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1524.  
XX  
XX Cytosstatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; Inflammatory; Immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406615P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406643P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406665P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476611P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 14-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang WM, Kochakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2755; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytosstatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 162 AA;  
  
Query Match 75.0%; Score 18; DB 8; Length 162;

Best Local Similarity 25.0%; Pred. No. 4e+02; Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
Db 53 CATTATTC 60

## RESULT 39

ID ADP30813 standard; protein, 162 AA.

AC ADP30813;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1580.

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN W0204035732-A2.

XX 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410961P.

PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.

PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.

PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.

PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.

PR 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411101P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467203P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485325P.

PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

## (FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang WM, Kotchakota S, Halshan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348436/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 2811; 428bp; English.

PS The present invention relates to an isolated nucleic acid molecule  
encoding a polypeptide which is believed to be cytosstatic, the  
antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.

XX SQ Sequence 162 AA;

Query Match 75.0%; Score 18; DB 8; Length 162;

Best Local Similarity 25.0%; Pred. No. 4e+02; Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
Db 107 CATTATTC 114

## RESULT 40

ID ADP31269 standard; protein, 165 AA.

AC ADP31269;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2036.

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

XX WO2004035732-A2.  
PN  
XX  
PD 29-APR-2004.  
PF 28-AUG-2003; 2003WO-US026780.  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410944P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Heetir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang WM, Kochakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as Proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3267; 428pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytosstatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.  
XX  
SQ Sequence 165 AA:  
  
Query Match 75.0%; Score 18; DB 8; Length 165;  
Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
Cy 1 CXXXXXXC 8  
Db 62 CTATTTTC 69  
  
RESULT 41  
ADP31099  
ID ADP31099 standard; protein; 168 AA.  
XX  
AC ADP31099;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1866.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
FN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.



PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410948P.  
 PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410958P.  
 PR 17-SEP-2002; 2002US-0410959P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410961P.  
 PR 17-SEP-2002; 2002US-0410962P.  
 PR 17-SEP-2002; 2002US-0411019P.  
 PR 17-SEP-2002; 2002US-0411022P.  
 PR 17-SEP-2002; 2002US-0411023P.  
 PR 17-SEP-2002; 2002US-0411024P.  
 PR 17-SEP-2002; 2002US-0411032P.  
 PR 17-SEP-2002; 2002US-0411035P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411041P.  
 PR 17-SEP-2002; 2002US-0411045P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411101P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471336P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493373P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Harenbeck RF, Huang MM, Kothakota S, Halsham L, Linemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 XX WPI; 2004-348438/32.  
 DR  
 XX  
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX  
 PS Claim 1; SEQ ID NO 3097; 428bp; English.  
 CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPWEB and is not in the specification.  
 XX  
 SQ Sequence 168 AA;  
 Query Match 75.0%; Score 18; DB 8; Length 168;  
 Best Local Similarity 25.0%; Pred. No. 4e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 CXXXXXXC 8  
 Db 25 CATTTTTC 32  
 RESULT 42  
 ADP30794  
 ID ADP30794 standard; protein; 171 AA.  
 AC ADP30794;  
 DT 12-AUG-2004 (first entry)  
 DE Human secreted protein SEQ ID #1561.  
 DE  
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KW cancer; inflammatory; immune; human secreted protein.  
 OS Homo sapiens.  
 PN WO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406642P.  
 PR 29-AUG-2002; 2002US-0406646P.  
 PR 29-AUG-2002; 2002US-0406653P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 29-AUG-2002; 2002US-0406666P.  
 PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410958P.  
 PR 17-SEP-2002; 2002US-0410959P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410961P.  
 PR 17-SEP-2002; 2002US-0410962P.  
 PR 17-SEP-2002; 2002US-0411019P.  
 PR 17-SEP-2002; 2002US-0411022P.  
 PR 17-SEP-2002; 2002US-0411023P.  
 PR 17-SEP-2002; 2002US-0411024P.  
 PR 17-SEP-2002; 2002US-0411032P.  
 PR 17-SEP-2002; 2002US-0411035P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411041P.  
 PR 17-SEP-2002; 2002US-0411045P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Hakenbeck RF, Huang MM, Kothakota S, Haisan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2792; 428pp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 171 AA;  
OY  
Db 1 CXXXXXXC 8  
161 CATATTC 168  
Query March 75 0%; Score 18; DB 8; Length 171;  
Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
RESULT 43  
ID ADP30793  
AC ADP30793 standard; protein; 171 AA.  
XX  
XX ADP30793;  
DT 12-AUG-2004 (first entry)

XX  
DE Human secreted protein SEQ ID #1560.  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Haislan L, Innemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
DR  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2791; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cyostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 171 AA:  
Query Match 75.0%; Score 18; DB 8; Length 171;  
Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Oy 1 CXXXXXXC 8  
Db 161 CATTATTC 168  
RESULT 44  
ID AAG10193 standard; protein; 177 AA.  
XX AAG10193;  
XX AAG10193;  
XX 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 8414.  
DE Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS EPI033405-A2.  
PN  
XX EPI033405-A2.  
PD  
XX 06-SEP-2000.  
PF  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-012180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0137222P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 15-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154033P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-015658P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.

PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 75.0%; Score 18; DB 3; Length 177;  
Best Local Similarity 25.0%; Pred. No. 4e+02; 6; Indels 0; Gaps 0;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
Db 87 CSSSSSC 94

## RESULT 45

ID ADP30755 standard; protein; 177 AA.

AC ADP30755;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1522.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

OS WO2004035732-A2.

PN 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

PF 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406640P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 17-SEP-2002; 2002US-0410946P.



PR 17-SEP-2002; 2002US-041105SP.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476009P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Heetir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang WM, Kochakota S, Halshan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3680; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytosstatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIOMBS and is not in the specification.  
XX  
SQ Sequence 177 AA;  
XX  
Query Match 75.0%; Score 18; DB 8; Length 177;  
Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
Db 39 CAAAAATC 46  
RESULT 47  
ADT56379  
ID ADT56379 standard; protein: 177 AA.  
XX  
XX AC ADT56379;  
XX  
DT 13-JAN-2005 (first entry)

XX  
DE Plant polypeptide, SEQ ID 6456.  
XX  
XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;  
XX disease resistance; galactomannan production; plant growth regulator;  
XX heat tolerance; herbicide tolerance; lignin production;  
XX extreme osmotic condition tolerance; pathogens resistance;  
XX pest resistance; yield improvement; seed oil yield; seed protein yield.  
XX  
OS Virdidplantae.  
XX  
XX US2004216190-A1.  
XX  
XX 28-OCT-2004.  
XX  
XX 18-DEC-2003; 2003US-00739930.  
XX  
XX 28-APR-2003; 2003US-00424599.  
XX  
XX 28-APR-2003; 2003US-00425115.  
XX  
XX (KOVA/) KOVALIC D K.  
XX  
XX KOVALIC DK;  
XX  
XX WPI; 2004-757369/74.  
XX  
XX  
XX New recombinant DNA constructs useful in the field of biochemistry and  
PT genetics, and in particular for producing transgenic plants with improved  
PT biological characteristics.  
XX  
XX Claim 2; SEQ ID NO 6456; 14dp; English.  
XX  
XX  
XX The invention relates a recombinant DNA construct comprising a  
CC polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:  
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences  
CC (SEQ ID NO: 5545-11088). The CDNAs and proteins are from corn, soybean,  
CC Arabidopsis, wheat and rape but the specification does not indicate which  
CC sequences is derived from which organism. Also included is a method of  
CC producing a plant having an improved property, comprising a transforming a  
CC plant with a recombinant DNA construct comprising a promoter region  
CC functional in a plant cell operably joined to a polynucleotide encoding a  
CC polypeptide associated with the property, and growing the transformed  
CC plant. The property is selected from improving plant cold tolerance, for  
CC manipulating growth rate in plant cells by modification of the cell cycle  
CC pathway, for improving plant drought tolerance, for providing increased  
CC resistance to plant disease, for galactomannan production, for production  
CC of plant growth regulators, for improving plant heat tolerance, for  
CC improving plant tolerance to herbicides, for increasing the rate of  
CC homologous recombination in plants, for lignin production, for improving  
CC plant tolerance to extreme osmotic conditions, for improving plant  
CC tolerance to pathogens or pests, for yield improvement by modification of  
CC photosynthesis, for modifying seed oil yield and/or content, for  
CC modifying seed protein yield and/or content, for yield improvement by  
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake  
CC and for yield improvement by providing improved plant growth and  
CC development under at least one stress condition. The polynucleotide may  
CC also encode a plant transcription factor. The methods and compositions of  
CC the present invention are useful in the field of biochemistry and  
CC genetics, in particular for producing transgenic plants with improved  
CC biological characteristics such as increased yield, improved nitrogen  
CC flow, increasing plant tolerance to cold or heat, improving plant  
CC tolerance to extreme osmotic and drought conditions, and improving plant  
CC tolerance to plant pests or pathogens. They can also be used in physical  
CC arrays of molecules, plant breeding markers, computer-based storage and  
CC analysis systems. The present sequence is one of the 5544 plant protein  
CC sequences of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.  
XX  
SQ Sequence 177 AA;  
XX  
Query Match 75.0%; Score 18; DB 8; Length 177;

Best Local Similarity 25.0%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
DB 87 CXXXXXXC 94

## RESULT 48

ABM86593  
ID ABM86593 standard; protein, 182 AA.

XX ABM86593;

XX 02-JUN-2005 (first entry)

DE Rice abiotic stress responsive polypeptide SEQ ID NO:4839.

KM abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.

XX Oryza sativa.

XX WO2003008540-A2.

XX 30-JAN-2003.

XX 21-JUN-2002; 2002WO-US019668.

XX 22-JUN-2001; 2001US-0300112P.

XX 24-AUG-2001; 2001US-0314662P.

XX 26-SEP-2001; 2001US-0325277P.

XX 21-NOV-2001; 2001US-0332132P.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Kregs J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;

XX Moughamer T, Provart N, Ricke D, Zhu T;

XX WPI; 2003-248011/24.

XX New stress-responsive nucleic acid, useful for altering the

XX responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold

XX stress, salt stress or osmotic stress.

XX Claim 1; SEQ ID NO 4839; 89pp; English.

XX The invention relates to novel abiotic stress responsive polynucleotides

XX and polypeptides. Also disclosed are vectors, expression cassettes, host

XX cells, and plants containing such polynucleotides. Also disclosed are

XX methods for using the polynucleotides and polypeptides to alter the

XX responsiveness of a plant to abiotic stress. The invention is useful in

XX agriculture. The nucleic acid is useful for determining whether a test

XX plant has been exposed to an abiotic stress condition. It is also useful

XX for selecting an agent that alters abiotic stress regulated

XX polynucleotide expression in a plant cell, and to identify a homolog or

XX ortholog to an abiotic stress responsive polynucleotide. The nucleic acid

XX molecule and the polypeptide encoded by it are useful in altering the

XX responsiveness of a plant to an abiotic stress, such as cold stress, salt

XX stress, osmotic stress or any of their combinations. The present sequence

XX is used in the exemplification of the invention

XX Sequence 182 AA;

XX Query Match 75.0%; Score 18; DB 7; Length 182;

XX Best Local Similarity 25.0%; Pred. No. 4e+02;

XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

XX 1 CXXXXXXC 8

XX DB 15 CSTASAC 22

## RESULT 49

ABM89559  
ID ABM89559 standard; protein, 185 AA.

XX ABM89559;

XX 02-JUN-2005 (first entry)

DE Rice abiotic stress responsive polypeptide SEQ ID NO:7805.

KM abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.

XX Oryza sativa.

XX WO2003008540-A2.

XX 30-JAN-2003.

XX 21-JUN-2002; 2002WO-US019668.

XX 22-JUN-2001; 2001US-0300112P.

XX 24-AUG-2001; 2001US-0314662P.

XX 26-SEP-2001; 2001US-0325277P.

XX 21-NOV-2001; 2001US-0332132P.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Kregs J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;

XX Moughamer T, Provart N, Ricke D, Zhu T;

XX WPI; 2003-248011/24.

XX New stress-responsive nucleic acid, useful for altering the

XX responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold

XX stress, salt stress or osmotic stress.

XX Claim 1; SEQ ID NO 7805; 89pp; English.

XX The invention relates to novel abiotic stress responsive polynucleotides

XX and polypeptides. Also disclosed are vectors, expression cassettes, host

XX cells, and plants containing such polynucleotides. Also disclosed are

XX methods for using the polynucleotides and polypeptides to alter the

XX responsiveness of a plant to abiotic stress. The invention is useful in

XX agriculture. The nucleic acid is useful for determining whether a test

XX plant has been exposed to an abiotic stress condition. It is also useful

XX for selecting an agent that alters abiotic stress regulated

XX polynucleotide expression in a plant cell, and to identify a homolog or

XX ortholog to an abiotic stress responsive polynucleotide. The nucleic acid

XX molecule and the polypeptide encoded by it are useful in altering the

XX responsiveness of a plant to an abiotic stress, such as cold stress, salt

XX stress, osmotic stress or any of their combinations. The present sequence

XX is used in the exemplification of the invention

XX Sequence 185 AA;

XX Query Match 75.0%; Score 18; DB 7; Length 185;

XX Best Local Similarity 25.0%; Pred. No. 4.1e+02;

XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

XX 1 CXXXXXXC 8

XX DB 148 CSTATTSC 155

## RESULT 50

ADP31576  
ID ADP31576 standard; protein, 192 AA.

XX ADP31576;

XX 12-AUG-2004 (first entry)

XX Human secreted protein SEQ ID #2343.

KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM cancer; inflammatory; immune; human secreted protein.  
XX Homo sapiens.  
XX MO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PE 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476099P.  
PR 09-JUN-2003; 2003US-047641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Halshan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
XX Claim 1; SEQ ID NO 3574; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytosratic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.  
XX  
XX Sequence 192 AA;  
SQ

Query Match 75.0%; Score 18; DB 8; Length 192;  
Best Local Similarity 25.0%; Pred. No. 4,1e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 CXXXXXXC 8  
Db 97 CTAAATC 104

Search completed: January 4, 2006, 15:56:14  
Job time : 56.7478 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time: 8.27626 seconds  
(without alignments)  
92.983 Million cell updates/sec

Title: US-09-932-322-9  
Perfect score: 24  
Sequence: 1 CXXXXXXC 8

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : PIR\_80:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	18	75.0	177 2 F96719	unknown protein, 3
2	18	75.0	355 2 A34597	rhamboid protein -
3	18	75.0	367 2 C39590	TPA-induced protei
4	18	75.0	492 2 S49147	ERF-2 protein - hu
5	18	75.0	843 1 JDLVJ1	DNA-directed DNA p
6	18	75.0	843 1 JDLVJ2	DNA-directed DNA p
7	18	75.0	1743 2 T26859	hypothetical prote
8	18	75.0	2543 2 T31687	surface antigen - p
9	17	70.8	26 2 A44036	collagen alpha 1(X
10	17	70.8	89 2 S72742	Bl177 F1.32 protei
11	17	70.8	101 2 T42262	hypothetical prote
12	17	70.8	120 2 S72979	probable L-a-amino
13	17	70.8	122 2 J00150	hypothetical 13k p
14	17	70.8	123 2 S25574	hypothetical 13k p
15	17	70.8	135 1 RWHUVY	T-cell receptor be
16	17	70.8	139 2 A71123	hypothetical prote
17	17	70.8	145 2 H87436	hypothetical prote
18	17	70.8	182 2 T23205	hypothetical prote
19	17	70.8	203 2 E70641	hypothetical prote
20	17	70.8	212 2 H75277	hypothetical prote
21	17	70.8	224 2 A41128	myogenin - human
22	17	70.8	230 2 S09778	hypothetical prote
23	17	70.8	237 2 P95411	ribonuclease (EC 3
24	17	70.8	238 1 JH0367	hypothetical prote
25	17	70.8	243 2 A35871	trypsin (EC 3.4.21
26	17	70.8	248 2 T47902	hypothetical prote
27	17	70.8	267 2 E87269	exopolysaccharide
28	17	70.8	271 2 F70771	probable glutamate
29	17	70.8	272 1 S72790	glutamate racemase

30	17	70.8	284 2	I51172
31	17	70.8	292 2	C64795
32	17	70.8	292 2	D90710
33	17	70.8	292 2	H85560
34	17	70.8	297 2	A98294
35	17	70.8	297 2	AH2989
36	17	70.8	305 2	AE0857
37	17	70.8	305 2	G65056
38	17	70.8	307 2	H85925
39	17	70.8	307 2	A91080
40	17	70.8	309 2	T32376
41	17	70.8	347 2	B53522
42	17	70.8	347 2	T36890
43	17	70.8	350 2	C84548
44	17	70.8	352 2	B87507
45	17	70.8	365 2	G71287
46	17	70.8	376 2	D97175
47	17	70.8	400 2	C34443
48	17	70.8	400 2	AE1988
49	17	70.8	408 1	WMNVIA
50	17	70.8	408 2	A72869
51	17	70.8	412 2	AG3637
52	17	70.8	419 2	C40901
53	17	70.8	429 2	T16656
54	17	70.8	439 2	A36385
55	17	70.8	456 2	S20597
56	17	70.8	461 2	A11858
57	17	70.8	467 2	JC6150
58	17	70.8	471 2	AS4840
59	17	70.8	471 2	S18446
60	17	70.8	489 2	S71955
61	17	70.8	492 1	WMAD55
62	17	70.8	539 2	H84640
63	17	70.8	541 2	T15299
64	17	70.8	544 2	S44814
65	17	70.8	547 2	S51475
66	17	70.8	567 2	T49942
67	17	70.8	577 2	B37057
68	17	70.8	615 2	B86212
69	17	70.8	631 2	T00925
70	17	70.8	638 2	D86477
71	17	70.8	655 2	AS9430
72	17	70.8	658 2	T19487
73	17	70.8	662 2	T17211
74	17	70.8	667 2	A48579
75	17	70.8	687 1	FYEFW
76	17	70.8	788 2	A37057
77	17	70.8	843 1	JDLVJ3
78	17	70.8	853 2	H70939
79	17	70.8	934 1	A34372
80	17	70.8	1028 2	T34360
81	17	70.8	1156 2	T23308
82	17	70.8	1279 2	A47363
83	17	70.8	1287 2	I46032
84	17	70.8	1345 2	S55659
85	17	70.8	1474 2	D88550
86	17	70.8	1522 2	H88380
87	17	70.8	1747 2	A45974
88	17	70.8	1857 2	S31212
89	17	70.8	1888 2	S50840
90	17	70.8	2395 1	S80820
91	17	70.8	2533 2	T28675
92	17	70.8	2533 2	T28674
93	17	70.8	2844 2	S28291
94	17	70.8	3124 2	A40020
95	17	70.8	3229 2	S27852
96	17	70.8	38 2	A47307
97	16	66.7	39 2	B41933
98	16	66.7	44 2	S38275
99	16	66.7	45 1	PFY2
100	16	66.7	55 1	NTNMB2
101	16	66.7	55 2	E82522
102	16	66.7	57 2	A46654

transcription fact  
c1g protein - Esc  
2-(5')-triphosphor  
2-(5')-triphosphor  
hypothetical prote  
permease (imported  
conserved hypotet  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
20k cytolophilin -  
probable iron-side  
hypothetical prote  
conserved hypotet  
nfs family enzyme  
nitrogenase cofact  
nitrogenase cofact  
immediate-early pr  
early gene transac  
cis,cis-muconate t  
hypothetical prote  
hypothetical prote  
surface antigen se  
ribonuclease-inhib  
ribonuclease-angio  
Gram-negative bact  
variant surface gl  
variant surface gl  
MDM2-like p53-bind  
early E1B 55K prot  
probable Rieseke ir  
hypothetical prote  
F4AB9.3 protein -  
cytochrome P450 cy  
hypothetical prote  
integrin beta-6 ch  
hypothetical prote  
hypothetical prote  
trophozoite surfac  
white protein - fr  
integrin beta-6 ch  
DNA-directed DNA p  
probable nitB prot  
complement C6 prec  
hypothetical prote  
hypothetical prote  
RNA helicase A - h  
nuclear DNA helica  
legumain protein 7  
protein ZC84.6 (im  
protein T22F7.3 (i  
collagen alpha 1(X  
collagen alpha 1(X  
collagen alpha 1(X  
surface protein ty  
alpha-51D immobili  
alpha-51D immobili  
hypothetical prote  
collagen alpha 1(X  
probable cell-surf  
heat-stable entero  
mating pheromone E  
toxin PLTX-II - sp  
fulvicin C - Myxoc  
neurotoxin B-II -  
hypothetical prote  
growth modulatory

103	16	66.7	57	2	B4654	growth modulatory
104	16	66.7	57	2	C4654	growth modulatory
105	16	66.7	57	2	S5436	metallothionein-2a
106	16	66.7	62	2	S5435	metallothionein-2c
107	16	66.7	72	1	WMBP2P	ogr protein - phag
108	16	66.7	72	1	DB2519	hypothetical prote
109	16	66.7	79	2	G82811	hypothetical prote
110	16	66.7	83	2	S11843	viral protein - Ag
111	16	66.7	83	2	AB3251	vira/g regulated p
112	16	66.7	91	2	G90824	hypothetical prote
113	16	66.7	92	2	D37057	epithelial cell gl
114	16	66.7	92	2	T30632	probable vitron co
115	16	66.7	100	2	A13337	hypothetical prote
116	16	66.7	104	2	B81045	hypothetical prote
117	16	66.7	109	2	S24253	Ig heavy chain V r
118	16	66.7	114	2	H71018	hypothetical prote
119	16	66.7	115	2	B96564	probable RING zinc
120	16	66.7	118	1	B5SNK1	phospholipase A2 (
121	16	66.7	118	1	B5SNK3	phospholipase A2 (
122	16	66.7	118	2	D34860	phospholipase A2 (
123	16	66.7	118	2	E34860	phospholipase A2 (
124	16	66.7	118	2	G34860	phospholipase A2 (
125	16	66.7	118	2	P34860	phospholipase A2 (
126	16	66.7	118	2	H34860	phospholipase A2 (
127	16	66.7	119	1	PSNOA1	phospholipase A2 h
128	16	66.7	119	2	G97003	hypothetical prote
129	16	66.7	120	2	E70424	plin - Aquifex ae
130	16	66.7	125	2	JH0184	hydrophobin Sc3 pr
131	16	66.7	126	2	D72784	hypothetical prote
132	16	66.7	132	2	I39193	gene HOXA1 protein
133	16	66.7	133	2	B30242	stem cell protein
134	16	66.7	135	2	D90847	hypothetical prote
135	16	66.7	136	2	I50554	gonadotropin II be
136	16	66.7	136	2	E71883	hypothetical prote
137	16	66.7	138	2	A80258	conserved hypotnet
138	16	66.7	140	2	T49481	hypothetical prote
139	16	66.7	140	2	T33359	hypothetical prote
140	16	66.7	141	2	T49380	hypothetical prote
141	16	66.7	141	2	PC1294	trichozoite surfac
142	16	66.7	142	2	J50510	fusaric acid resis
143	16	66.7	143	1	QOEC32	hypothetical prote
144	16	66.7	144	1	PSHUPP	phospholipase A2 (
145	16	66.7	144	2	JN0480	phospholipase A2 (
146	16	66.7	144	2	S01391	phospholipase A2 (
147	16	66.7	144	2	A60480	interleukin-9 prec
148	16	66.7	145	1	PSKFXU	phospholipase A2 (
149	16	66.7	145	2	S07983	phospholipase A2 h
150	16	66.7	145	2	S01390	phospholipase A2 (
151	16	66.7	145	2	S74292	hypothetical prote
152	16	66.7	145	2	T42280	hypothetical prote
153	16	66.7	145	2	T15608	hypothetical prote
154	16	66.7	148	1	PSHU	phospholipase A2 (
155	16	66.7	149	2	C81959	probable pilin NMA
156	16	66.7	150	2	C42744	exonuclease (EC 3.
157	16	66.7	151	2	T25047	hypothetical prote
158	16	66.7	153	2	S34255	hypothetical prote
159	16	66.7	157	2	S55956	probable membrane
160	16	66.7	160	2	J00542	185k secretory pro
161	16	66.7	166	2	P84449	hypothetical prote
162	16	66.7	172	2	AB3514	outer membrane pro
163	16	66.7	173	2	T14827	lipid transfer pro
164	16	66.7	174	2	AC0306	probable outer mem
165	16	66.7	177	2	G86917	conserved hypotnet
166	16	66.7	177	2	D70654	hypothetical prote
167	16	66.7	178	2	G71195	hypothetical prote
168	16	66.7	181	2	JN0342	N-methyl-D-asparta
169	16	66.7	181	2	T15412	knutz type subcil
170	16	66.7	183	2	JX0311	knutz type subcil
171	16	66.7	183	2	I38923	microfibril-associ
172	16	66.7	183	2	A54151	microfibril-associ
173	16	66.7	183	2	A49313	microfibril-associ
174	16	66.7	183	2	S55599	hypothetical prote
175	16	66.7	186	2	E82625	outer membrane pro
176	16	66.7	187	2	S36918	zygote-specific pr
177	16	66.7	187	2	P90315	hypothetical prote
178	16	66.7	188	2	T33886	hypothetical prote
179	16	66.7	189	2	G70347	phenylacrylic acid
180	16	66.7	190	2	JX0310	knutz type subcil
181	16	66.7	190	2	A59416	knutz type subcil
182	16	66.7	190	2	A59418	knutz type subcil
183	16	66.7	195	2	B82857	PilX protein XE003
184	16	66.7	198	2	F83435	conserved hypotnet
185	16	66.7	199	2	S35441	serine-tRNA ligase
186	16	66.7	201	2	S23324	gene achate prote
187	16	66.7	202	2	AC2875	hypothetical prote
188	16	66.7	204	2	T28939	hypothetical prote
189	16	66.7	205	2	T37947	hypothetical prote
190	16	66.7	207	2	B81418	hypothetical prote
191	16	66.7	208	2	S28675	hypothetical prote
192	16	66.7	208	2	P00278	hypothetical prote
193	16	66.7	208	2	B64925	probable oxidoredu
194	16	66.7	208	2	B90926	probable oxidoredu
195	16	66.7	208	2	A85775	probable oxidoredu
196	16	66.7	213	2	F75585	urease accessory p
197	16	66.7	220	1	S71282	ribosomal protein
198	16	66.7	220	1	T21730	hypothetical prote
199	16	66.7	222	2	I39192	hypothetical prote
200	16	66.7	223	2	A65172	gene HOXA1 protein
201	16	66.7	224	2	A36675	myogenin - mouse
202	16	66.7	225	2	B65127	type 4 prepilin-11
203	16	66.7	225	2	T49769	hypothetical prote
204	16	66.7	225	2	T33361	hypothetical prote
205	16	66.7	227	2	B35882	myogenin - chicken
206	16	66.7	229	1	TRBOTR	trypsin (EC 3.4.21
207	16	66.7	229	1	T29203	hypothetical prote
208	16	66.7	231	1	TRPGRR	trypsin (EC 3.4.21
209	16	66.7	231	2	T02765	glutathione transf
210	16	66.7	233	2	S71464	ribonuclease (EC 3
211	16	66.7	234	2	T52018	fumarate reductase
212	16	66.7	237	2	S08073	cyclic nucleotide
213	16	66.7	237	2	T25152	hypothetical prote
214	16	66.7	238	1	W2BE2	gene 2 protein - h
215	16	66.7	238	2	S31779	trypsin (EC 3.4.21
216	16	66.7	240	2	S39047	trypsin (EC 3.4.21
217	16	66.7	240	2	H82289	conserved hypotnet
218	16	66.7	240	2	T25641	hypothetical prote
219	16	66.7	240	2	A43912	myogenin - Japanes
220	16	66.7	241	2	S39048	trypsin (EC 3.4.21
221	16	66.7	242	2	C70895	hypothetical prote
222	16	66.7	242	2	T45795	hypothetical prote
223	16	66.7	245	2	T24565	hypothetical prote
224	16	66.7	247	2	S13813	trypsin (EC 3.4.21
225	16	66.7	247	2	S05494	trypsin (EC 3.4.21
226	16	66.7	248	2	S55067	trypsin (EC 3.4.21
227	16	66.7	248	2	T21786	hypothetical prote
228	16	66.7	248	2	F85067	hypothetical prote
229	16	66.7	249	2	T32060	hypothetical prote
230	16	66.7	254	2	G71545	probable deoxyoctu
231	16	66.7	254	2	A56447	CMP-2-xeto-3-deoxy
232	16	66.7	258	2	T32542	hypothetical prote
233	16	66.7	259	2	T46881	succinate dehydrog
234	16	66.7	261	2	F82728	succinate dehydrog
235	16	66.7	262	2	F72003	probable outer mem
236	16	66.7	262	2	G86619	probable omp limpo
237	16	66.7	262	2	F87253	dienelactone hydro
238	16	66.7	262	2	A57391	TSL-2 protein pre
239	16	66.7	265	2	A95401	protein limported
240	16	66.7	265	2	S06462	beta-lactamase (EC
241	16	66.7	266	2	S33163	beta-lactamase (EC
242	16	66.7	267	2	S22684	beta-lactamase (EC
243	16	66.7	268	2	T10304	inhibitor of apopt
244	16	66.7	268	2	A53989	apoptosis-inhibiti
245	16	66.7	272	1	A36082	insulin-like growt
246	16	66.7	274	2	T47933	hypothetical prote
247	16	66.7	276	2	T09131	chitinase (EC 3.2.
248	16	66.7	279	2	C70458	diaminopimelate ep

249	16	66.7	280	2	T09504	322	16	66.7	366	2	A95382	probable iron upta
250	16	66.7	283	2	A64174	323	16	66.7	367	2	T49048	hypothetical prote
251	16	66.7	287	2	T15779	324	16	66.7	369	2	AE2345	phospho-N-acetylmu
252	16	66.7	288	2	T04401	325	16	66.7	369	2	G81101	membrane-bound lyt
253	16	66.7	288	2	T21790	326	16	66.7	369	2	T15213	hypothetical prote
254	16	66.7	288	2	T18075	327	16	66.7	370	2	B64632	hydrogenase expres
255	16	66.7	289	1	Q2AD2	328	16	66.7	370	2	F71883	hydrogenase expres
256	16	66.7	289	1	Q2AD5	329	16	66.7	371	2	A39234	opsin - bluebotle
257	16	66.7	292	2	T23966	330	16	66.7	372	2	C97791	nifs protein homol
258	16	66.7	293	2	T03883	331	16	66.7	373	1	OOFF	rhodopsin - fruit
259	16	66.7	294	2	T23682	332	16	66.7	374	2	D83328	probable aminotran
260	16	66.7	296	2	S77427	333	16	66.7	374	2	S40691	opsin rh1 - fruit
261	16	66.7	297	2	A10827	334	16	66.7	374	2	T05660	hypothetical prote
262	16	66.7	297	2	T38161	335	16	66.7	374	2	T31842	zinc finger protei
263	16	66.7	298	2	AF0578	336	16	66.7	376	2	B16522	iron-sulfur cofact
264	16	66.7	302	2	B71310	337	16	66.7	377	2	C89456	protein F55A4.8 [i
265	16	66.7	303	2	T19289	338	16	66.7	378	2	A82406	cytochrome d ubiq
266	16	66.7	304	2	B83052	339	16	66.7	379	2	S67856	Gung proteome d ubiq
267	16	66.7	305	2	T49720	340	16	66.7	379	2	F71533	hypothetical prote
268	16	66.7	307	1	GSFF3	341	16	66.7	379	2	A81678	conserved hypotnet
269	16	66.7	309	2	S41427	342	16	66.7	381	2	A35300	G protein-coupled
270	16	66.7	309	2	S37028	343	16	66.7	383	2	I53870	Ekg-1 orphan recep
271	16	66.7	309	2	S39957	344	16	66.7	384	2	S68410	lipase Arab-1 - Ar
272	16	66.7	309	2	C59977	345	16	66.7	384	2	F97711	hypothetical prote
273	16	66.7	311	2	S20085	346	16	66.7	384	2	T49084	hypothetical prote
274	16	66.7	312	1	QOBE07	347	16	66.7	385	2	T04725	hypothetical prote
275	16	66.7	313	2	D85855	348	16	66.7	385	2	T06484	aspartate carboxy
276	16	66.7	313	2	B91011	349	16	66.7	387	2	C41880	nitrogenase cofact
277	16	66.7	313	2	B49885	350	16	66.7	388	2	T16861	hypothetical prote
278	16	66.7	314	2	T27686	351	16	66.7	389	2	T23167	hypothetical prote
279	16	66.7	314	2	B43652	352	16	66.7	391	1	J01626	actachment protein
280	16	66.7	318	2	B91177	353	16	66.7	391	2	T08338	conserved hypotnet
281	16	66.7	319	2	JC2252	354	16	66.7	392	2	B47071	nitrogenase cofact
282	16	66.7	321	2	T02236	355	16	66.7	395	2	F82657	beta-ketoacyl-[ACP
283	16	66.7	322	2	S54806	356	16	66.7	399	2	T21015	hypothetical prote
284	16	66.7	323	1	A48149	357	16	66.7	404	2	A46274	HIV gp120-binding
285	16	66.7	323	2	T03614	358	16	66.7	411	2	E97338	3-oxoacyl-(acyl)-ca
286	16	66.7	323	2	S56777	359	16	66.7	411	2	S58105	Cu metalloregulato
287	16	66.7	328	2	C72370	360	16	66.7	416	2	S27198	homeoic protein H
288	16	66.7	328	2	T36494	361	16	66.7	418	2	T15142	hypothetical prote
289	16	66.7	329	2	UC7935	362	16	66.7	424	2	JC5891	omega 6 desaturase
290	16	66.7	329	2	AB1327	363	16	66.7	425	2	B90113	seryl-tRNA synthet
291	16	66.7	329	2	AD1698	364	16	66.7	425	2	JC5909	AS3 protein - fru
292	16	66.7	330	1	UN0561	365	16	66.7	425	2	A37912	thrombin receptor
293	16	66.7	330	2	I56100	366	16	66.7	427	2	G71272	iron-sulfur cofact
294	16	66.7	330	2	C86023	367	16	66.7	429	2	T21113	hypothetical prote
295	16	66.7	330	2	T19644	368	16	66.7	431	2	T36682	probable secreted
296	16	66.7	331	2	T26004	369	16	66.7	432	2	C70381	large subunit of i
297	16	66.7	331	2	A30242	370	16	66.7	432	2	A43448	thrombin receptor
298	16	66.7	332	2	AD3535	371	16	66.7	435	2	T15143	hypothetical prote
299	16	66.7	334	2	H69214	372	16	66.7	439	2	T28841	hypothetical prote
300	16	66.7	335	2	G01448	373	16	66.7	442	2	T24196	hypothetical prote
301	16	66.7	335	2	B71361	374	16	66.7	442	2	D84600	probable xyuliose
302	16	66.7	337	2	A53041	375	16	66.7	444	2	F86329	hypothetical prote
303	16	66.7	338	2	C68986	376	16	66.7	448	2	S41725	integrase - Saccha
304	16	66.7	338	2	T06603	377	16	66.7	449	2	T26571	hypothetical prote
305	16	66.7	340	2	T48288	378	16	66.7	450	2	T17234	hypothetical prote
306	16	66.7	344	1	RWRTC2	379	16	66.7	450	2	AD3117	polylacturonase
307	16	66.7	344	2	G82428	380	16	66.7	453	2	A98170	hypothetical prote
308	16	66.7	344	2	S72723	381	16	66.7	456	2	A31857	ribonuclease inhib
309	16	66.7	345	2	B85881	382	16	66.7	456	2	T37575	bacterial cell cyc
310	16	66.7	345	2	B65012	383	16	66.7	466	2	AD3477	malate dehydrogena
311	16	66.7	345	2	H91036	384	16	66.7	470	2	A40697	63K sperm flagella
312	16	66.7	345	2	A05279	385	16	66.7	474	2	S31712	beta-1,3-glucanase
313	16	66.7	347	2	S41638	386	16	66.7	475	2	S53389	glutamate decarbox
314	16	66.7	348	2	JQ0431	387	16	66.7	475	2	T50686	peroxisomal Ca-dep
315	16	66.7	349	2	S53864	388	16	66.7	477	1	I38893	transcription fact
316	16	66.7	349	2	T16882	389	16	66.7	478	2	S31906	beta-1,3-glucanase
317	16	66.7	355	2	C96651	390	16	66.7	481	2	T20309	hypothetical prote
318	16	66.7	355	2	T00578	391	16	66.7	486	2	JC7212	stretch-activated
319	16	66.7	358	2	D89007	392	16	66.7	486	2	T38087	hypothetical prote
320	16	66.7	359	2	T26813	393	16	66.7	487	2	F84727	hypothetical prote
321	16	66.7	364	2	B36313	394	16	66.7	488	1	JC2385	protein disulfide-

395	16	66.7	492	2	JCS169	alkaline nuclease
396	16	66.7	496	2	D83614	conserved hypobeth
397	16	66.7	498	2	AP0157	probable aldehyde
398	16	66.7	498	2	A48203	interleukin-14 pre
399	16	66.7	499	2	B81914	probable periplasm
400	16	66.7	504	1	A28807	protein disulfide-
401	16	66.7	504	1	S41661	protein disulfide-
402	16	66.7	504	2	G02474	interferon regulat
403	16	66.7	505	1	B69417	probable thymidine
404	16	66.7	505	1	SS5507	protein disulfide-
405	16	66.7	505	1	SS3994	protein disulfide-
406	16	66.7	505	2	SS8363	protein disulfide-
407	16	66.7	505	2	JCS704	hydroxymethylgluta
408	16	66.7	509	2	S46314	carboxypeptidase-1
409	16	66.7	509	2	H70597	Rtn-like protein l
410	16	66.7	510	2	T48977	probable zinc meta
411	16	66.7	510	2	H90787	serine-trna ligase
412	16	66.7	512	2	T37819	dopamine receptor-
413	16	66.7	514	2	G01026	hypothetical prote
414	16	66.7	514	2	D56849	phosphatidylinosit
415	16	66.7	519	2	T45764	hypothetical prote
416	16	66.7	521	2	A09100	probable interleuk
417	16	66.7	521	2	T27192	hypothetical prote
418	16	66.7	522	2	S71821	probable hsp-60 -
419	16	66.7	526	2	T13687	60 kda chaperonin
420	16	66.7	528	2	T33599	hypothetical prote
421	16	66.7	533	2	H71492	probable hsp-60 -
422	16	66.7	534	2	B81654	60 kda chaperonin
423	16	66.7	536	2	H85647	hypothetical prote
424	16	66.7	547	1	A32244	60k cysteine-rich
425	16	66.7	547	2	S19607	alkaline phosphata
426	16	66.7	547	2	B43584	60k cysteine-rich
427	16	66.7	548	2	B84306	hypothetical prote
428	16	66.7	553	2	D71515	60k cysteine-rich
429	16	66.7	554	2	C81671	60 kda outer membr
430	16	66.7	554	2	T25288	hypothetical prote
431	16	66.7	556	1	S12602	60k cysteine-rich
432	16	66.7	556	2	A86560	60 kda cysteine-ri
433	16	66.7	557	1	B39439	60k cysteine-rich
434	16	66.7	557	2	D84496	probable retroelem
435	16	66.7	557	2	A48434	valiant-specific s
436	16	66.7	558	2	JCS204	60k cysteine-rich
437	16	66.7	561	2	S63388	probable membrane
438	16	66.7	569	2	D85354	hypothetical prote
439	16	66.7	573	2	T49610	related to rna bin
440	16	66.7	577	2	T28017	hypothetical prote
441	16	66.7	579	2	T24880	hypothetical prote
442	16	66.7	589	2	T43210	fibulin-ID precurs
443	16	66.7	591	2	T39195	probable amino aci
444	16	66.7	600	2	B46642	DNA-directed DNA p
445	16	66.7	601	2	B96744	unknown protein [i
446	16	66.7	604	2	T49577	hypothetical prote
447	16	66.7	617	2	T19444	luciferin A related
448	16	66.7	631	2	S71508	dnak-lype molecula
449	16	66.7	638	2	G02068	white homolog - hu
450	16	66.7	643	2	T25473	hypothetical prote
451	16	66.7	655	2	T51792	hypothetical prote
452	16	66.7	655	2	JCS2005	integrin beta-5 ch
453	16	66.7	657	2	B72486	hypothetical prote
454	16	66.7	668	2	A42275	chemoreceptor prot
455	16	66.7	672	2	T12524	hypothetical prote
456	16	66.7	677	2	C42125	trophozoite cystei
457	16	66.7	680	2	PN0510	integrin beta-3 ch
458	16	66.7	686	2	S30075	ferric reductase (
459	16	66.7	686	2	S43562	KOBE5.3 proteain -
460	16	66.7	689	2	T42760	fibulin, splice fo
461	16	66.7	692	2	T32980	hypothetical prote
462	16	66.7	704	2	S21911	BRCore-NS-Z3 prote
463	16	66.7	704	2	T31227	tran protein homol
464	16	66.7	706	2	A81848	probable Tonb-depe
465	16	66.7	707	2	G86894	hypothetical prote
466	16	66.7	709	2	T29692	hypothetical prote
467	16	66.7	712	2	T42990	fibulin 1, splice
468	16	66.7	713	2	I50128	fibroblast growth
469	16	66.7	713	2	A35502	major surface-labe
470	16	66.7	722	2	E71403	hypothetical prote
471	16	66.7	723	2	PN0509	integrin beta-3 ch
472	16	66.7	734	2	UC4861	feritilin beta cha
473	16	66.7	735	2	G02937	hypothetical prote
474	16	66.7	743	2	T02147	SNF1-related prote
475	16	66.7	746	2	S62365	hypothetical prote
476	16	66.7	752	2	T48574	hypothetical prote
477	16	66.7	752	2	T20871	platelet glycoprot
478	16	66.7	753	2	B36268	platelet glycoprot
479	16	66.7	758	2	T48815	mixed-linked glucu
480	16	66.7	761	2	A46193	88k B-26-specific
481	16	66.7	763	2	E96693	probable terpene s
482	16	66.7	767	2	T07693	hypothetical prote
483	16	66.7	769	2	S50966	probable membrane
484	16	66.7	770	2	I40027	nitrile reductase
485	16	66.7	773	2	I46059	beta-1 integrin su
486	16	66.7	778	2	A60798	platelet glycoprot
487	16	66.7	781	2	S43534	integrin beta3 - c
488	16	66.7	788	2	A26547	platelet glycoprot
489	16	66.7	788	2	T77349	platelet glycoprot
490	16	66.7	788	2	I51530	integrin beta-3 su
491	16	66.7	798	2	B27079	fibronectin recept
492	16	66.7	798	2	A28193	integrin beta-1 ch
493	16	66.7	798	2	B28193	integrin beta-1* c
494	16	66.7	798	2	T22793	hypothetical prote
495	16	66.7	799	2	A38308	integrin beta-5 ch
496	16	66.7	801	2	G82302	probable cellulob
497	16	66.7	805	2	I40029	nitrile reductase
498	16	66.7	807	2	E83726	assimilatory nitri
499	16	66.7	810	2	B30848	plasma (EC 3.4.21
500	16	66.7	817	2	T25674	hypothetical prote
501	16	66.7	818	2	JC4058	fibroblast growth
502	16	66.7	822	2	S19947	fibroblast growth
503	16	66.7	822	2	B49151	fibroblast growth
504	16	66.7	822	2	T01622	probable salt-indu
505	16	66.7	823	2	D86165	protein f15K9.3 [i
506	16	66.7	829	2	T45683	fibroblast growth
507	16	66.7	834	2	T09369	hypothetical prote
508	16	66.7	838	2	T04449	hypothetical prote
509	16	66.7	849	2	E86306	similar to tuftell
510	16	66.7	849	2	T01286	probable RNA-bind
511	16	66.7	849	2	S51789	VDL receptor prec
512	16	66.7	863	1	S33015	hypothetical prote
513	16	66.7	872	2	E75489	conserved hypobeth
514	16	66.7	885	2	JN0339	N-methyl-D-asparta
515	16	66.7	889	2	D71414	hypothetical prote
516	16	66.7	897	2	S67283	hypothetical prote
517	16	66.7	901	2	JN0337	N-methyl-D-asparta
518	16	66.7	905	2	T23229	hypothetical prote
519	16	66.7	906	2	JN0341	N-methyl-D-asparta
520	16	66.7	906	2	A46286	N-methyl-D-asparta
521	16	66.7	907	2	E96692	probable wall-abso
522	16	66.7	913	2	G64110	hypothetical prote
523	16	66.7	922	2	JN0340	N-methyl-D-asparta
524	16	66.7	922	2	JN0338	N-methyl-D-asparta
525	16	66.7	935	2	S64384	probable membrane
526	16	66.7	938	2	A46612	N-methyl-D-asparta
527	16	66.7	938	2	S21104	N-methyl-D-asparta
528	16	66.7	938	2	S19710	N-methyl-D-asparta
529	16	66.7	943	2	A47551	N-methyl-D-asparta
530	16	66.7	946	2	T01460	hypothetical prote
531	16	66.7	959	2	JN0336	N-methyl-D-asparta
532	16	66.7	964	2	S48404	probable membrane
533	16	66.7	964	1	F1BVCV	RNA 1 protein - cu
534	16	66.7	992	2	A31666	hypothetical prote
535	16	66.7	1007	2	T01437	hypothetical prote
536	16	66.7	1025	2	B54718	dihydropyrimidine
537	16	66.7	1025	2	A54718	probable membrane
538	16	66.7	1029	2	S56229	nitrile reductase
539	16	66.7	1044	2	T43155	nitrile reductase
540	16	66.7	1044	2	T43160	nitrile reductase

541	16	66.7	1046	2	A26838	614	15	62.5	27	2	PC4234	hypothetical prote
542	16	66.7	1069	2	S27922	615	15	62.5	31	2	B61014	defensin R-2 - rat
543	16	66.7	1084	2	T15616	616	15	62.5	34	2	C44336	neurotoxin Tx-3 -
544	16	66.7	1099	2	T16283	617	15	62.5	35	2	S18224	filamentous hemag
545	16	66.7	1101	2	T16840	618	15	62.5	40	2	UJ0515	Ig heavy chain V-I
546	16	66.7	1129	2	A47511	619	15	62.5	42	2	A17907	lytrotropin beta c
547	16	66.7	1132	2	JW0091	620	15	62.5	45	2	AC3539	hypothetical prote
548	16	66.7	1132	2	JC4127	621	15	62.5	46	2	B81072	hypothetical prote
549	16	66.7	1153	2	P84468	622	15	62.5	46	2	CS3613	plectoxin IX - spi
550	16	66.7	1218	2	B84537	623	15	62.5	52	1	XBPI	bromelain inhibito
551	16	66.7	1235	2	CE9165	624	15	62.5	52	2	S66609	bromelain inhibito
552	16	66.7	1236	2	T50904	625	15	62.5	52	2	T10299	conotoxin-like pro
553	16	66.7	1274	2	T10729	626	15	62.5	53	2	T10405	conotoxin-like pro
554	16	66.7	1297	2	T52065	627	15	62.5	53	2	C72850	conotoxin homolog
555	16	66.7	1297	2	T30274	628	15	62.5	53	2	T30499	conotoxin-like pro
556	16	66.7	1307	2	G96711	629	15	62.5	54	1	FEME	ferredoxin 214Fe-4
557	16	66.7	1316	2	G70535	630	15	62.5	56	1	TRPG	pancreatic secreto
558	16	66.7	1321	2	JE0352	631	15	62.5	57	1	NILSA	main precursor -
559	16	66.7	1336	2	T23310	632	15	62.5	58	2	D82759	hypothetical prote
560	16	66.7	1347	2	T41321	633	15	62.5	59	2	T41417	B3 protein [import
561	16	66.7	1357	2	T16860	634	15	62.5	59	2	AE2319	hypothetical prote
562	16	66.7	1360	2	T33922	635	15	62.5	62	2	F82332	bacterioferritin-a
563	16	66.7	1373	2	JE0095	636	15	62.5	63	2	S25772	testis-specific pr
564	16	66.7	1376	2	VCBE40	637	15	62.5	63	2	A34484	metallothionein I
565	16	66.7	1396	1	S63986	638	15	62.5	64	2	AD2881	hypothetical prote
566	16	66.7	1413	2	T46354	639	15	62.5	66	2	T17934	hypothetical prote
567	16	66.7	1416	2	B88550	640	15	62.5	68	2	B84730	hypothetical prote
568	16	66.7	1416	2	B88550	641	15	62.5	69	2	B97230	hypothetical prote
569	16	66.7	1444	2	T18856	642	15	62.5	70	2	S23316	hypothetical prote
570	16	66.7	1469	2	B36665	643	15	62.5	70	2	CS3620	hypothetical prote
571	16	66.7	1475	2	A60026	644	15	62.5	71	2	F69981	hypothetical prote
572	16	66.7	1509	2	T19486	645	15	62.5	71	2	T00008	copy number contro
573	16	66.7	1518	2	T28880	646	15	62.5	71	2	A59412	KCP-bearing platel
574	16	66.7	1523	2	T13935	647	15	62.5	71	2	A59413	platelet-aggregati
575	16	66.7	1531	2	T42218	648	15	62.5	71	2	S13168	baroxostatin ba
576	16	66.7	1557	2	T28811	649	15	62.5	71	2	G43019	platelet aggregati
577	16	66.7	1559	2	T16210	650	15	62.5	72	2	A42325	platelet aggregati
578	16	66.7	1620	2	T27283	651	15	62.5	72	2	AB2528	platelet aggregati
579	16	66.7	1650	2	S53457	652	15	62.5	72	2	F43019	platelet aggregati
580	16	66.7	1700	2	S08167	653	15	62.5	72	2	D43019	platelet aggregati
581	16	66.7	1772	2	A45532	654	15	62.5	72	2	T43019	platelet aggregati
582	16	66.7	1784	2	T02844	655	15	62.5	72	2	B43019	platelet aggregati
583	16	66.7	1801	1	MMRTS	656	15	62.5	73	2	B43020	platelet aggregati
584	16	66.7	1822	2	S63985	657	15	62.5	73	2	T42302	hypothetical prote
585	16	66.7	1896	2	T01490	658	15	62.5	73	2	S78719	protein YER091c-a
586	16	66.7	2022	2	T48818	659	15	62.5	73	2	H43019	platelet aggregati
587	16	66.7	2049	2	T29227	660	15	62.5	73	2	C43019	platelet aggregati
588	16	66.7	2180	2	T29764	661	15	62.5	73	2	E43019	platelet aggregati
589	16	66.7	2182	2	T14320	662	15	62.5	73	2	A43020	platelet aggregati
590	16	66.7	2233	2	T28669	663	15	62.5	73	2	A43019	platelet aggregati
591	16	66.7	2287	2	T21312	664	15	62.5	73	2	B40003	platelet aggregati
592	16	66.7	2399	2	H71879	665	15	62.5	73	2	A40003	platelet aggregati
593	16	66.7	2470	2	I50726	666	15	62.5	74	2	S33833	hypothetical prote
594	16	66.7	2529	2	B64635	667	15	62.5	74	2	H90820	hypothetical prote
595	16	66.7	2531	2	T16743	668	15	62.5	75	2	S19696	mating pheromone B
596	16	66.7	2548	2	B59435	669	15	62.5	76	1	TRAM3	proteinase inhibit
597	16	66.7	2643	2	T28149	670	15	62.5	76	2	I83048	FSH beta-subunit -
598	16	66.7	2704	2	S09118	671	15	62.5	76	2	G59226	hypothetical prote
599	16	66.7	2718	2	A23475	672	15	62.5	76	2	A39034	neurogranin - bovi
600	16	66.7	2761	2	T21064	673	15	62.5	76	2	D44007	apoptoxin IV - tra
601	16	66.7	3133	2	S52093	674	15	62.5	76	2	C44007	apoptoxin VI - tra
602	16	66.7	3672	2	T23433	675	15	62.5	76	2	AH2120	hypothetical prote
603	16	66.7	3704	2	T37316	676	15	62.5	77	2	JN0536	hypothetical 8.3K
604	16	66.7	4056	2	H86599	677	15	62.5	78	2	D84620	hypothetical prote
605	16	66.7	4544	1	S02339	678	15	62.5	78	2	G98261	hypothetical prote
606	16	66.7	4545	1	S25111	679	15	62.5	79	2	C60767	retrovirus-related
607	16	66.7	4548	1	S00657	680	15	62.5	79	2	AH3369	hypothetical prote
608	16	66.7	4660	2	T42737	681	15	62.5	80	2	I65235	testicular luteini
609	16	66.7	4753	1	A47437	682	15	62.5	81	2	CE5019	hypothetical prote
610	16	66.7	5107	2	T29144	683	15	62.5	81	2	AC1864	hypothetical prote
611	16	66.7	10797	2	T30192	684	15	62.5	83	1	W8BP67	gene 18.7 protein
612	15	62.5	15	2	B59045	685	15	62.5	83	2	F35962	hypothetical prote
613	15	62.5	27	2	S55030	686	15	62.5	83	2	A34156	blisteratin - puff

687	15	62.5	85	1	W9BPB7	gene 19.2 protein	760	15	62.5	119	2	PH1554	Ig H chain V regio
688	15	62.5	85	2	A1Z765	hypothetical prote	761	15	62.5	119	2	A45574	13.3 kda polypepti
689	15	62.5	86	2	S20471	class V zygote-ape	762	15	62.5	120	1	PSKFP3	phospholipase A2 (
690	15	62.5	87	2	T18115	hypothetical prote	763	15	62.5	120	1	PSKFP3	phospholipase A2 (
691	15	62.5	88	2	P88129	protein P5281.6 (i	764	15	62.5	120	2	I50993	gonadotropin I bet
692	15	62.5	89	2	P90453	ferredoxin like pr	765	15	62.5	120	2	G71105	hypothetical prote
693	15	62.5	91	2	C38180	hypothetical prote	766	15	62.5	120	2	A12181	hypothetical prote
694	15	62.5	92	2	B90856	hypothetical prote	767	15	62.5	121	1	PSBGAC	phospholipase A2 (
695	15	62.5	94	2	T03285	anther-specific pr	768	15	62.5	121	1	RWBEM3	surface glycoprote
696	15	62.5	95	1	BORT3	prostatic steroid-	769	15	62.5	121	2	PH1661	Ig heavy chain V r
697	15	62.5	96	2	A57483	3-mercaptopyruvate	770	15	62.5	122	1	PSABA	phospholipase A2 (
698	15	62.5	97	2	B97832	alpha-(1,3)-fucosy	771	15	62.5	122	1	PSRSAM	phospholipase A2 (
699	15	62.5	98	2	T33302	hypothetical prote	772	15	62.5	122	1	PSRSAM	phospholipase A2 (
700	15	62.5	98	2	S42596	hypothetical prote	773	15	62.5	122	1	PSVII	phospholipase A2 i
701	15	62.5	100	2	G87427	hypothetical prote	774	15	62.5	122	2	A53872	phospholipase A2 (
702	15	62.5	101	1	WMVZ11	11k protein - vacc	775	15	62.5	122	2	A25806	phospholipase A2 (
703	15	62.5	101	1	MMVZ12	11k protein - vacc	776	15	62.5	122	2	A44179	phospholipase A2 (
704	15	62.5	101	1	T30793	probable DNA-bindi	777	15	62.5	122	2	S46979	phospholipase A2 -
705	15	62.5	101	2	S73700	probable lipoprote	778	15	62.5	122	2	S62780	phospholipase A2 (
706	15	62.5	102	2	S50530	hypothetical prote	779	15	62.5	122	2	JX0063	phospholipase A2 (
707	15	62.5	102	2	S69884	hypothetical prote	780	15	62.5	122	2	S09314	phospholipase A2 (
708	15	62.5	102	2	C61889	hypothetical prote	781	15	62.5	122	2	T49039	response regulator
709	15	62.5	102	2	B82751	hypothetical prote	782	15	62.5	123	2	B53872	phospholipase A2 (
710	15	62.5	103	2	C69000	hypothetical prote	783	15	62.5	124	1	PSPGA2	phospholipase A2 (
711	15	62.5	103	2	I58936	hypothetical prote	784	15	62.5	124	1	NRCM	pancreatic ribonuc
712	15	62.5	104	2	D71035	hypothetical prote	785	15	62.5	124	1	NRCM	pancreatic ribonuc
713	15	62.5	104	2	B81283	hypothetical prote	786	15	62.5	124	1	NRCM	pancreatic ribonuc
714	15	62.5	104	2	S50911	metallothionein-2	787	15	62.5	124	2	A26535	phospholipase A2 (
715	15	62.5	106	2	S74084	folilitropin beta c	788	15	62.5	124	2	A59420	phospholipase A2 (
716	15	62.5	106	2	B69048	conserved hypotnet	789	15	62.5	124	2	I51190	phospholipase A2 -
717	15	62.5	106	2	S32032	Sp17 protein precu	790	15	62.5	124	2	S25113	insulin-like growt
718	15	62.5	108	2	T45063	hypothetical prote	791	15	62.5	125	2	JX0052	phospholipase A2 (
719	15	62.5	109	2	PH1653	Ig heavy chain V r	792	15	62.5	125	2	S65624	probable membrane
720	15	62.5	109	2	T46171	hypothetical prote	793	15	62.5	125	2	AD0297	hypothetical prote
721	15	62.5	109	2	B70956	hypothetical prote	794	15	62.5	125	2	T28743	CS9 protein - bab
722	15	62.5	110	2	S55208	hypothetical prote	795	15	62.5	126	2	I36914	hypothetical prote
723	15	62.5	111	1	A29654	proteinase inhibit	796	15	62.5	126	2	F84240	conserved hypotnet
724	15	62.5	111	2	S69911	Ig V-D-J region (R	797	15	62.5	126	2	G90350	Killer toxin KP4 P
725	15	62.5	111	2	H91187	hypothetical prote	798	15	62.5	127	2	S40034	Rabl protein homo
726	15	62.5	111	2	H85046	hypothetical prote	799	15	62.5	127	2	T10546	surface glycoprote
727	15	62.5	112	2	B29654	gonadotropin beta	800	15	62.5	128	1	RMHU59	CS9 protein - gre
728	15	62.5	113	1	S07092	hypothetical prote	801	15	62.5	128	2	I36894	probable phosphori
729	15	62.5	113	2	T36789	probable membrane	802	15	62.5	128	2	T35073	hypothetical membr
730	15	62.5	115	2	S66924	phosphoribosyl-AMP	803	15	62.5	128	2	AH3422	folilitropin beta c
731	15	62.5	115	2	T45252	phospholipase A2 (	804	15	62.5	129	1	FTTHUB	folilitropin beta c
732	15	62.5	117	2	A34860	phospholipase A2 (	805	15	62.5	129	1	FTPCB	folilitropin beta c
733	15	62.5	117	2	B32416	phospholipase A2 (	806	15	62.5	129	2	A23550	folilitropin beta c
734	15	62.5	117	2	A32416	phospholipase A2 (	807	15	62.5	129	2	S73514	probable lipoprote
735	15	62.5	117	2	T26924	hypothetical prote	808	15	62.5	130	2	JC4526	folilitropin beta c
736	15	62.5	118	1	PSKFT1	phospholipase A2 (	809	15	62.5	130	2	A32893	complement regulat
737	15	62.5	118	1	PSKFT2	phospholipase A2 (	810	15	62.5	131	2	I56894	pigment deposition
738	15	62.5	118	1	PSKFT3	phospholipase A2 (	811	15	62.5	131	2	A46298	hypothetical prote
739	15	62.5	118	1	PSOXB	folilitropin beta c	812	15	62.5	131	2	A83350	phospholipase A2 (
740	15	62.5	118	1	FTTHOB	folilitropin beta c	813	15	62.5	132	1	PSHOA	agouti protein pre
741	15	62.5	118	2	A32622	phospholipase A2 (	814	15	62.5	132	2	I37143	hypothetical prote
742	15	62.5	118	2	B34860	phospholipase A2 (	815	15	62.5	132	2	B83382	tryptsin inhibitor
743	15	62.5	118	2	C34860	phospholipase A2 (	816	15	62.5	133	1	PSOXG	textilotoxin chain
744	15	62.5	118	2	S01801	phospholipase A2 (	817	15	62.5	133	2	T1RZBR	hypothetical prote
745	15	62.5	118	2	PN0141	lutropin beta chai	818	15	62.5	133	2	S14728	hypothetical prote
746	15	62.5	118	2	PN0139	lutropin beta chai	819	15	62.5	134	2	H84868	hypothetical prote
747	15	62.5	118	2	S61051	hypothetical prote	820	15	62.5	134	1	C64872	hypothetical prote
748	15	62.5	118	2	T49515	hypothetical prote	821	15	62.5	135	2	E83461	hypothetical prote
749	15	62.5	118	2	T32397	hypothetical prote	822	15	62.5	135	2	T15610	hypothetical prote
750	15	62.5	119	1	PSNJ2K	phospholipase A2 (	823	15	62.5	136	2	AB0846	formate hydrogenly
751	15	62.5	119	1	PSNJ3K	phospholipase A2 (	824	15	62.5	136	2	S08626	formate hydrogenly
752	15	62.5	119	1	PSBYA	phospholipase A2 (	825	15	62.5	136	2	F91075	formate hydrogenly
753	15	62.5	119	1	PSNJ2W	phospholipase A2 (	826	15	62.5	136	2	H82787	hypothetical prote
754	15	62.5	119	1	PSNJ3W	phospholipase A2 (	827	15	62.5	136	2	T23996	hypothetical prote
755	15	62.5	119	1	PSOXA	phospholipase A2 (	828	15	62.5	136	2	T36241	hypothetical prote
756	15	62.5	119	2	JN0403	phospholipase A2 (	829	15	62.5	137	1	JC4877	phospholipase A2 h
757	15	62.5	119	2	S07528	phospholipase A2 (	830	15	62.5	137	2	JC4710	beta-bungarotoxin
758	15	62.5	119	2	UX0115	pancreatic ribonuc	831	15	62.5	137	2	S37848	hypothetical prote
759	15	62.5	119	2	A61465	lutropin beta chai	832	15	62.5	137	2	T49243	hypothetical prote



979 15 62.5 174 2 A86358 Similar to blue co  
980 15 62.5 175 2 T48541 hypothetical prote  
981 15 62.5 176 2 T18921 hypothetical prote  
982 15 62.5 177 1 AFMD8 phycoerythrin beta  
983 15 62.5 177 1 A55537 CDdiacylglycerol-  
984 15 62.5 177 2 C71329 hypothetical prote  
985 15 62.5 178 2 T29345 hypothetical prote  
986 15 62.5 178 2 C84723 probable PHD-type  
987 15 62.5 178 2 D75390 NADH dehydrogenase  
988 15 62.5 180 2 S14019 hypothetical prote  
989 15 62.5 180 2 B84768 hypothetical prote  
990 15 62.5 180 2 F75607 arsenate reductase  
991 15 62.5 181 2 D86518 hypothetical prote  
992 15 62.5 181 2 C72104 hypothetical prote  
993 15 62.5 182 2 T45302 hypothetical prote  
994 15 62.5 182 2 T11906 NADH2 dehydrogenas  
995 15 62.5 183 2 A85561 hypothetical prote  
996 15 62.5 183 2 D64795 Apo-citrate lyase  
997 15 62.5 183 2 E90710 Apo-citrate lyase  
998 15 62.5 184 2 AG0736 hypothetical prote  
999 15 62.5 185 1 JC2394 phospholipase A2 1  
1000 15 62.5 185 2 C96808 protein F28K19.4 [

## ALIGNMENTS

## RESULT 1

Unknown protein, 30275-30808 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C/Accession: F96719

R/Theologian: A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Hulzar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marzilli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: F96719  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-177 <STO>  
A/Cross-references: UNIPROT:Q9C9L7; UNIPARC:UPI00000A1912; GB:AEO05173; NID:9665540; P C/Genetics:  
A/Map position: 1

Query Match 75.0%; Score 18; DB 2; Length 177;  
Best Local Similarity 25.0%; Pred. No. 66;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
Db 87 CSSSSSSC 94

## RESULT 2

rhomboid protein - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 06-Jul-1990 #sequence\_revision 06-Jul-1990 #text\_change 05-Oct-2004

C/Accession: A34597

R/Bier, E.; Jan, L.Y.; Jan, Y.N.  
Genes Dev. 4, 190-203, 1990

A/Title: rhomboid, a gene required for dorsoventral axis establishment and peripheral ne  
A/Reference number: A34597; MUID:90249726; PMID:2110920  
A/Accession: A34597

A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-355 <BIE>  
A/Cross-references: UNIPROT:P20350; UNIPARC:UPI000016BD27; GB:X52454; NID:910878; PIDN:C C/Genetics:  
A/Map position: 1

Query Match 75.0%; Score 18; DB 2; Length 355;  
Best Local Similarity 25.0%; Pred. No. 82;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
Db 57 CSTASSC 64

## RESULT 3

TPA-induced protein 11D - mouse

C/Species: Mus musculus (house mouse)

C/Date: 30-Aug-1991 #sequence\_revision 30-Aug-1991 #text\_change 09-Jul-2004

C/Accession: C39590

R/Varnum, B.C.; Ma, O.; Chi, T.; Fletcher, B.; Herschman, H.R.  
Mol. Cell. Biol. 11, 1754-1758, 1991  
A/Title: The TIS11 primary response gene is a member of a gene family that encodes prote

A/Reference number: A39590; MUID:91141531; PMID:1956120

A/Accession: C39590

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-367 <VAR>

A/Cross-references: UNIPROT:P23949; UNIPARC:UPI00002A130; GB:M58564; NID:9202066; PIDN:C C/Keywords: DNA binding; zinc finger

Query Match 75.0%; Score 18; DB 2; Length 367;  
Best Local Similarity 25.0%; Pred. No. 82;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
Db 273 CSSSASSC 280

## RESULT 4

ERF-2 protein - human

C/Species: Homo sapiens (man)

C/Date: 16-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004

C/Accession: S49147

R/Nie, X.F.; Maclean, R.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.  
submitted to the EMBL Data Library, April 1994

A/Reference number: S49147

A/Accession: S49147

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-492 <NIE>

A/Cross-references: UNIPROT:P47974; UNIPARC:UPI000003794E  
R/Nie, X.F.; Maclean, R.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.  
Gene 152, 285-286, 1995

A/Title: ERF-2, the human homologue of the murine Tis11d early response gene.

A/Reference number: 137372; MUID:95137407; PMID:7835719

A/Accession: 137372

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-492 <RES>

A/Cross-references: UNIPARC:UPI000003794E; EMBL:X78992; NID:9509777; PIDN:CAA55592.1; PIR C/Genetics:

A/Map position: 13q12.3-13q12.3

A/Reference number: 13q12.3-13q12.3



Query Match 75.0%; Score 18; DB 2; Length 492;  
Best Local Similarity 25.0%; Pred. No. 90;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
DB 299 CASSSSSC 306

## RESULT 5

JDVLU1  
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adw, strain Japan/  
C/Species: hepatitis B virus, HBV  
C/Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 09-Jul-2004  
C/Accession: D28925  
R/Okamoto, H.; Tsuda, F.; Sakugawa, H.; Sastrosoewignjo, R.I.; Imai, M.; Miyakawa, Y.; M  
J. Gen. Virol. 69, 2575-2583, 1988  
A/Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of sur  
A/Reference number: J50253; MUID:89010694; PMID:3171552  
A/Accession: D28925  
A/Molecule type: DNA  
A/Residues: 1-843 <OKA>  
A/Cross-references: UNIPROT:P17394; UNIPARC:UPI00001297F9; GB:D00329; NID:G221497  
C/Superfamily: hepatitis virus DNA-directed DNA polymerase  
C/Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 75.0%; Score 18; DB 1; Length 843;  
Best Local Similarity 25.0%; Pred. No. 1.1e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
DB 268 CASSSSSC 275

## RESULT 6

JDVLU2  
DNA-directed DNA polymerase (EC 2.7.7.7) - hepatitis B virus (subtype adw, strain Okinaw  
C/Species: hepatitis B virus, HBV  
C/Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 09-Jul-2004  
C/Accession: E28925  
R/Okamoto, H.; Tsuda, F.; Sakugawa, H.; Sastrosoewignjo, R.I.; Imai, M.; Miyakawa, Y.; M  
J. Gen. Virol. 69, 2575-2583, 1988  
A/Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of sur  
A/Reference number: J50253; MUID:89010694; PMID:3171552  
A/Accession: E28925  
A/Molecule type: DNA  
A/Residues: 1-843 <OKA>  
A/Cross-references: UNIPROT:P17395; UNIPARC:UPI00001297FC; GB:D00330; NID:G221498  
C/Superfamily: hepatitis virus DNA-directed DNA polymerase  
C/Keywords: DNA biosynthesis; nucleotidyltransferase

Query Match 75.0%; Score 18; DB 1; Length 843;  
Best Local Similarity 25.0%; Pred. No. 1.1e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
DB 268 CASSSSSC 275

## RESULT 7

T26859  
hypothetical protein Y43F8B.3 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T26859  
R/Ainscough, R.  
submitted to the EMBL Data Library, October 1998  
A/Accession: T26859  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA

A/Residues: 1-1743 <NIL>  
A/Cross-references: UNIPROT:Q9XWX5; UNIPARC:UPI0000078FD6; EMBL:AL032623; PIDN:CAA21511.  
A/Experimental source: clone Y43F8B  
C/Genetics:

A/Gene: CESP:Y43F8B.3  
A/Introns: 65/1; 92/2; 128/1; 229/1; 367/1; 422/1; 486/2; 523/1; 571/1; 628/1; 857/2; 94

Query Match 75.0%; Score 18; DB 2; Length 1743;  
Best Local Similarity 25.0%; Pred. No. 1.3e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
DB 1123 CSATATSTC 1130

## RESULT 8

T31687  
surface antigen - Paramoecium primaurelia  
C/Species: Paramoecium primaurelia  
C/Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C/Accession: T31687  
R/Bourgain-Guglielmi, F.; Caron, F.  
Journal of Eukaryot. Microbiol. 43, 303-314, 1996  
A/Title: Molecular characterization of the D surface protein gene subfamily in Paramoeci  
A/Reference number: 221061; MUID:96313351; PMID:8768434  
A/Accession: T31687  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-2543 <BOU>  
A/Cross-references: UNIPROT:P90649; UNIPARC:UPI0000079A03; EMBL:X96616; NID:G1235576; PI  
C/Genetics:  
A/Genetic code: SGC5  
C/Superfamily: G surface protein

Query Match 75.0%; Score 18; DB 2; Length 2543;  
Best Local Similarity 25.0%; Pred. No. 1.5e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
DB 1961 CTSSTTTC 1968

## RESULT 9

A44036  
collagen alpha 1(XII) chain - bovine (fragment)  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C/Accession: A44036  
R/Instrum, G.P.; McDonough, A.M.; Martinovitch, M.P.; Keene, D.R.; Morris, N.P.; Burgeso  
J. Biol. Chem. 267, 20087-20092, 1992  
A/Title: Identification and partial purification of a large, variant form of type XII co  
A/Reference number: A44036; MUID:93015874; PMID:1400326  
A/Accession: A44036  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-26 <LUN>  
A/Cross-references: UNIPROT:Q9TRM6; UNIPARC:UPI0000087EAB  
A/Experimental source: skin  
A/Note: sequence extracted from NCBI backbone (NCBIP:115795)

Query Match 70.8%; Score 17; DB 2; Length 26;  
Best Local Similarity 25.0%; Pred. No. 89;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
DB 10 CETATSTC 17

## RESULT 10

S72742

B1177\_F1\_32 protein - Mycobacterium leprae  
C/Species: Mycobacterium leprae  
C/Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C/Accession: S72742  
R/Smith, D.R.; Robison, K.  
submitted to the EMBL Data Library, November 1993  
A/Description: Mycobacterium leprae cosmid B1177.  
A/Reference number: S72694  
A/Accession: S72742  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-89 <SM1>  
A/Cross-references: UNIPROT:Q49652; UNIPARC:UPI00000B6610; EMBL:U00011; NID:g466807; PID  
C/Genetics:  
A/Start codon: GTG

Query Match 70.8%; Score 17; DB 2; Length 89;  
Best Local Similarity 25.0%; Pred. No. 1.3e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
Db 71 CRTSSTAC 78

RESULT 11  
T42262  
hypothetical protein - phage SPPI  
C/Species: phage SPPI  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: T42262  
R/Alonso, J.C.; Iudert, G.; Stiege, A.C.; Chal, S.; Weise, F.; Trautner, T.A.  
Gene 204, 201-212, 1997  
A/Title: The complete nucleotide sequence and functional organization of Bacillus subtilis  
A/Reference number: Z22137; MUID:98094274; PMID:9434185  
A/Accession: T42262  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-101 <ALO>  
A/Cross-references: UNIPROT:O48435; UNIPARC:UPI000009B544; EMBL:X97918; PIDN:CAA66572.1

Query Match 70.8%; Score 17; DB 2; Length 101;  
Best Local Similarity 25.0%; Pred. No. 1.3e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
Db 59 CSKTSSAC 66

RESULT 12  
S72979  
probable L-a-aminoadipyl-L-cystenyl-D-valine synthetase acvS - Mycobacterium leprae  
N/Alternate names: B229\_C2\_179 protein  
C/Species: Mycobacterium leprae  
C/Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C/Accession: S72979  
R/Smith, D.R.; Robison, K.  
submitted to the EMBL Data Library, November 1993  
A/Description: Mycobacterium leprae cosmid B229.  
A/Reference number: S72588  
A/Accession: S72979  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-120 <SM1>  
A/Cross-references: UNIPROT:Q49859; UNIPARC:UPI00000BE2D7; EMBL:U00020; NID:g467102; PID  
C/Genetics:  
A/Gene: acvS  
A/Start codon: GTG

Query Match 70.8%; Score 17; DB 2; Length 120;  
Best Local Similarity 25.0%; Pred. No. 1.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
Db 4 CSYTSASC 11

RESULT 13  
J00150  
hypothetical 13k protein - Pseudomonas aeruginosa  
C/Species: Pseudomonas aeruginosa  
C/Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Sep-1999  
C/Accession: J00150  
R/Kaco, J.; Chu, L.; Kitano, K.; DeVault, J.D.; Kimbara, K.; Chakrabarty, A.M.; Misra, T.  
Gene 84, 31-38, 1989  
A/Title: Nucleotide sequence of a regulatory region controlling alginate synthesis in P.  
A/Reference number: J00132; MUID:90108714; PMID:2514124  
A/Accession: J00150  
A/Status: translation not shown  
A/Molecule type: DNA  
A/Residues: 1-122 <KAT>  
A/Cross-references: UNIPARC:UPI00001791D0  
C/Superfamily: Pseudomonas aeruginosa hypothetical 13k protein

Query Match 70.8%; Score 17; DB 2; Length 122;  
Best Local Similarity 25.0%; Pred. No. 1.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
Db 49 CSTATVAC 56

RESULT 14  
S25574  
Ig heavy chain V region - human  
C/Species: Homo sapiens (man)  
C/Date: 25-Feb-1994 #sequence\_revision 26-May-1995 #text\_change 20-Jun-2000  
C/Accession: S25574  
R/Adderson, E.B.; Shackelford, P.G.; Quinn, A.; Carroll, W.L.  
J. Immunol. 147, 1667-1674, 1991  
A/Title: Restrictcd Ig H chain V gene usage in the human antibody response to Haemophilus  
A/Reference number: S25571; MUID:9134592; PMID:1908880  
A/Accession: S25574  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-123 <ADD>  
A/Cross-references: UNIPARC:UPI0000115EBF; EMBL:X56527; NID:G33569; PIDN:CAA39874.1; PID  
A/Note: the authors translated the codon GAA for residue 69 as Asp and TTA for residue 10  
C/Superfamily: immunoglobulin V region; immunoglobulin homology  
C/Keywords: heterotetramer; immunoglobulin  
F/23-108/Domain: immunoglobulin homology <IMW>

Query Match 70.8%; Score 17; DB 2; Length 123;  
Best Local Similarity 25.0%; Pred. No. 1.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
Db 113 CASSAYAC 120

RESULT 15  
RKHUY  
T-cell receptor beta chain precursor V region (YT35) - human  
N/Alternate names: T-cell receptor beta-1 chain J-B1.2 segment  
C/Species: Homo sapiens (man)  
C/Date: 03-Aug-1984 #sequence\_revision 03-Aug-1984 #text\_change 09-Jul-2004  
C/Accession: A02000; E24687  
R/Yanagi, Y.; Yoshikai, Y.; Leggett, K.; Clark, S.P.; Aleksander, I.; Mak, T.W.  
Nature 308, 145-149, 1984  
A/Title: A human T cell-specific cDNA clone encodes a protein having extensive homology to  
A/Reference number: A93324; MUID:84142269; PMID:6336315  
A/Accession: A02000

A:Molecule type: mRNA  
 A:Residues: 1-135 <YAN>  
 A:Cross-references: UNIPROT:P01733; UNIPARC:UPI00001377D3; GB:K01571  
 A:Experimental source: clone YT35  
 R:Toyomaga, B.; Yoshikawa, Y.; Vadasz, V.; Chin, B.; Mak, T. W.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 8624-8628, 1985  
 A:Title: Organization and sequences of the diversity, joining, and constant region genes  
 A:Reference number: A94081; MUID:86094276; PMID:3866244  
 A:Accession: E24687  
 A:Molecule type: DNA  
 A:Residues: 121-135 <TOY>  
 A:Cross-references: UNIPARC:UPI0000113C76; GB:M4159; NID:9338844; PIDN:AAA60669.1; PID:  
 C:Genetics:  
 A:Gene: GDB:TCRB  
 A:Cross-references: GDB:120405; OMIM:186930  
 A:Map position: 7q35-7q35  
 C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: glycoprotein; heterotetramer; receptor; T-cell  
 F:1-21/Domain: signal sequence #status predicted <Sig>  
 F:22-135/Product: T-cell receptor beta chain V region YT35 #status predicted <MAT>  
 F:35-113/Domain: immunoglobulin homology <IMM>  
 F:42-111/Disulfide bonds: #status predicted  
 F:90/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 70.8%; Score 17; DB 1; Length 135;  
 Best Local Similarity 25.0%; Pred. No. 1.5e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
 Db 111 CASSFTC 118

RESULT 16  
 A71123  
 hypothetical protein PH0752 - *Pyrococcus horikoshii*  
 C:Species: *Pyrococcus horikoshii*  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 12-Jul-2004  
 C:Accession: A71123  
 R:Kawabata, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
 M.; Ohnuki, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
 DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: A71123  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-139 <KAW>  
 A:Cross-references: UNIPROT:O58508; UNIPARC:UPI0000628CE; GB:AP000003; NID:93236130; PI  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH0752

Query Match 70.8%; Score 17; DB 2; Length 139;  
 Best Local Similarity 25.0%; Pred. No. 1.5e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
 Db 81 CSSTYASC 88

RESULT 17  
 H87436  
 hypothetical protein CC1513 [imported] - *Caulobacter crescentus*  
 C:Species: *Caulobacter crescentus*  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
 C:Accession: H87436  
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Land, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
 u, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: H87436  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-145 <STO>  
 A:Cross-references: UNIPROT:Q9A851; UNIPARC:UPI00000C73F0; GB:AE005673; NID:g13422892; P  
 C:Genetics:  
 A:Gene: CC1513

Query Match 70.8%; Score 17; DB 2; Length 145;  
 Best Local Similarity 25.0%; Pred. No. 1.5e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
 Db 10 CAABEAC 17

RESULT 18  
 T22305  
 hypothetical protein F46C5.2 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T22305  
 R:Thomas, K.  
 submitted to the EMBL Data Library, September 1995  
 A:Reference number: Z19544  
 A:Accession: T22305  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-182 <WIL>  
 A:Cross-references: UNIPROT:P52881; UNIPARC:UPI000013A06E; EMBL:Z54281; PIDN:CAA91050.1;  
 C:Genetics:  
 A:Experimental source: clone F46C5  
 A:Gene: CESP:F46C5.2  
 A:Map position: 2  
 A:Introns: 14/3; 50/1; 76/3; 145/2

Query Match 70.8%; Score 17; DB 2; Length 182;  
 Best Local Similarity 25.0%; Pred. No. 1.6e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
 Db 35 CTSGATTTC 42

RESULT 19  
 E70641  
 hypothetical protein Rv0698 - *Mycobacterium tuberculosis* (strain H37RV)  
 C:Species: *Mycobacterium tuberculosis*  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
 C:Accession: E70641  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garlier, T.; Churcher, C.; Harris, D.; Gordon, S.  
 Comor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
 Rajandram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Suleiron, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: E70641  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-203 <COL>  
 A:Cross-references: UNIPROT:P95044; UNIPARC:UPI00000C14B5; GB:Z84395; GB:AL123456; NID:g  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: Rv0698

Query Match 70.8%; Score 17; DB 2; Length 203;  
 Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
Db 51 CTTAAHTC 58

## RESULT 20

H75277  
hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: H75277

R:White, O.; Eissen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma  
Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: H75277

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-212 <WHI>

A:Cross-references: UNIPROT:Q9RKS8; UNIPARC:UPI00000CIAF7; GB:AE002071; GB:AE000513; NID

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2407

A:Map position: 1

Query Match 70.8%; Score 17; DB 2; Length 212;  
Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
Db 15 CRSASAC 22

## RESULT 21

A41128  
myogenin - human

N:Alternate names: myogenic determination factor Myf-4

C:Species: Homo sapiens (man)

C:Date: 03-Aug-1992 #sequence\_revision 03-Aug-1992 #text\_change 09-Jul-2004

C:Accession: A41128; S27295; S06948; S21978

R:Salminen, A.; Braun, T.; Buchberger, A.; Uuers, S.; Winter, B.; Arnold, H.H.

J. Cell Biol. 115, 905-917, 1991

A:Title: Transcription of the muscle regulatory gene MYF4 is regulated by serum component

A:Reference number: A41128; MUID:9204650; PMID:1659574

A:Accession: A41128

A:Molecule type: DNA

A:Residues: 1-224 <SLA>

A:Cross-references: UNIPROT:P15173; UNIPARC:UPI000012FBA4; EMBL:X62155

R:Braun, T.; Bober, E.; Buschhausen-Denker, G.; Kohz, S.; Grzeschik, K.H.; Arnold, H.H.

EMBO J. 9, 592, 1990

A:Reference number: S27295

A:Contents: extratum

A:Accession: S27295

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-224 <BRN>

A:Cross-references: UNIPARC:UPI000012FBA4

A:Note: this is a revision to the sequence from reference S06947

R:Braun, T.; Bober, E.; Buschhausen-Denker, G.; Kotz, S.; Grzeschik, K.H.; Arnold, H.H.

EMBO J. 8, 3617-3625, 1989

A:Title: Differential expression of myogenic determination genes in muscle cells: possit

A:Reference number: S06947; MUID:90059960; PMID:2583111

A:Accession: S06948

A:Molecule type: mRNA

A:Residues: 1-154; 'SOGCPANALATAPAVOSGAVHWSAPTOGIISRLTLQWPTCTSPSPSWTASQWKCGLMPQMKPC

A:Cross-references: UNIPARC:UPI000016AD95; EMBL:X17651; NID:934831; PIDN:CAA35641.1; PID

A:Note: this sequence has been revised in reference S27295

R:Arnold, H.H.

submitted to the EMBL Data Library, September 1991

A:Reference number: S21978

A:Accession: S21978

A:Molecule type: DNA

A:Residues: 1-223 <ARN>

A:Cross-references: UNIPARC:UPI000016AD96; EMBL:X62155; NID:934833; PIDN:CAA44080.1; PID

C:Genetics:

A:Gene: GDB:MYOG; MYF4

A:Cross-references: GDB:120210; OMIM:159980

A:Map position: 1q31-1q41

A:Introns: 157/3; 184/3

C:Superfamily: human myogenin

C:Keywords: DNA binding

Query Match 70.8%; Score 17; DB 2; Length 224;  
Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
Db 162 CSHSASC 169

## RESULT 22

S09778  
hypothetical protein UL16 precursor - human cytomegalovirus (strain AD169)

C:Species: human cytomegalovirus, human herpesvirus 5

A:Note: host Homo sapiens (man)

C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004

C:Accession: S09778

R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; F

M.; Barrett, B.G.

Curr. Top. Microbiol. Immunol. 154, 125-169, 1990

A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus

A:Reference number: S09749; MUID:90269039; PMID:2161319

A:Accession: S09778

A:Molecule type: DNA

A:Residues: 1-230 <CHE>

A:Cross-references: UNIPROT:P16757; UNIPARC:UPI0000137B56; EMBL:X17403; NID:959591; PIDN

A:Note: this sequence was submitted to the EMBL Data Library, December 1989

C:Superfamily: human cytomegalovirus hypothetical protein UL16

C:Keywords: glycoprotein; transmembrane protein

F:1-55/Domain: signal sequence #status predicted <SIG>

F:26-230/Product: hypothetical protein UL16 #status predicted <MAT>

F:189-206/Domain: transmembrane #status predicted <TM>

F:35;41;68;84;95;101;132;145/Binding site: darbyhydrate (asn) (covalent) #status predicted

Query Match 70.8%; Score 17; DB 2; Length 230;  
Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
Db 18 CLSASSC 25

## RESULT 23

F95411  
hypothetical protein Sma222 [imported] - Sinorhizobium meliloti (strain 1021) megaplasm

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004

C:Accession: F95411

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: F95411

A:Molecule type: preliminary

A:Residues: 1-237 <KR>

A:Cross-references: UNIPROT:Q92XP6; UNIPARC:UPI00000CB344; GB:AE006469; PIDN:AAK5856.1;

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; LaLaurie, M.; Vandenbol, M.; Vorkhler, F.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.  
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: Sma2221  
A:Genome: plasmid

Query Match 70.8%; Score 17; DB 2; Length 237;  
Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 76 CTATTTC 83

RESULT 24  
JH0367  
ribonuclease (EC 3.1.27.-) Rh precursor - *Rhizopus niveus*  
N:Alternate names: base nonspecific ribonuclease  
C:Species: *Rhizopus niveus*  
C:Date: 05-Mar-1993 #sequence\_revision 31-Jan-1997 #text\_change 31-Dec-2004  
C:Accession: JH0367; A1449; B4149  
C:Obj1, K.; Horluchi, H.; Watanabe, H.; Takagi, M.; Yano, K.; Irie, M.  
J. Biochem. 109, 776-785, 1991  
A:Title: Expression of RNase Rh from *Rhizopus niveus* in yeast and characterization of the  
A:Reference number: JH0367; MUID:92011465; PMID:1655721  
A:Accession: JH0367  
A:Molecule type: DNA  
A:Residues: 1-238 <OHG>  
A:Cross-references: UNIPROT:P08056; UNIPARC:UPI00001344AC; DBJ:D12476; DBJ:D01125; NID  
J. Horluchi, H.; Yano, K.; Takagi, M.; Yano, K.; Wakabayashi, E.; Sada, A.; Mine, S.; C  
J. Biochem. 109, 408-418, 1988  
A:Title: Primary structure of a base non-specific ribonuclease from *Rhizopus niveus*.  
A:Reference number: A1449; MUID:88273061; PMID:3391995  
A:Accession: A1449  
A:Molecule type: DNA  
A:Residues: 1-81, 'SLV', 85-238 <HOR>  
A:Cross-references: UNIPARC:UPI000016897A; GB:D00238; NID:G218044; PIDN:BA00167.1; PID  
A:Note: the authors translated the codon TCA for residue 82 as Asn, CTC for residue 83  
A:Accession: B4149  
A:Molecule type: protein  
A:Residues: 17-238 <HO2>  
R:Kuribara, H.; Nomura, T.; Mitsui, Y.; Ohgi, K.; Irie, M.; Nakamura, K.T.  
J. Mol. Biol. 255, 310-320, 1996  
A:Title: The crystal structure of ribonuclease Rh from *Rhizopus niveus* at 2.0 Å resolution  
A:Reference number: S62310; MUID:96150305; PMID:8551522  
A:Contents: annotation; X-ray crystallography, 2.0 angstroms  
C:Genetics:  
A:Introns: 21/1; 116/2; 151/3; 167/2  
A:Function:  
A:Description: hydrolyzes internal phosphodiester bonds of RNA to produce 3'-phosphomono  
C:Superfamily: RNases  
C:Keywords: hydrolase  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-238/Product: ribonuclease Rh #status experimental <MAT>  
F:19-36, 26-69, 75-136, 79-128, 198-229/Disulfide bonds: #status experimental  
F:62, 121, 125/Active site: His, Glu, His #status predicted  
F:65, 120, 124/Binding site: substrate (Trp, His, Lys) #status predicted

Query Match 70.8%; Score 17; DB 1; Length 238;  
Best Local Similarity 25.0%; Pred. No. 1.7e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8

Db 19 CSSTALSC 26

RESULT 25  
A5871  
trypsin (EC 3.4.21.4) precursor, pancreatic - African clawed frog  
C:Species: *Xenopus laevis* (African clawed frog)  
C:Date: 09-Nov-1990 #sequence\_revision 09-Nov-1990 #text\_change 09-Jul-2004  
C:Accession: A5871; S12117  
R:Shi, Y.B.; Brown, D.D.  
Genes Dev. 4, 1107-1113, 1990  
A:Title: Developmental and thyroid hormone-dependent regulation of pancreatic genes in *X*  
A:Reference number: A5871; MUID:91007255; PMID:2210372  
A:Accession: A5871  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-243 <SH>  
A:Cross-references: UNIPROT:P19799; UNIPARC:UPI00001376E0; EMBL:X53458; NID:965162; PIDN  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; protein digestion; serine proteinase  
F:1-15/Domain: signal sequence #status predicted <SIG>  
F:16-20/Domain: activation peptide #status predicted <APT>  
F:21-236/Domain: trypsin homology <TRY>  
F:26-243/Product: trypsin I #status predicted <MAT>  
F:27-157, 45-61, 129-230, 168-182/Disulfide bonds: #status predicted  
F:60, 104, 197/Active site: His, Asp, Ser #status predicted

Query Match 70.8%; Score 17; DB 2; Length 243;  
Best Local Similarity 25.0%; Pred. No. 1.8e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 129 CSAAGTSC 136

RESULT 26  
T47902  
hypotheical protein T4C21.320 - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T47902  
R:Choisme, N.; Robert, C.; Brothier, P.; Catolico, L.; Artiguenave, F.; San  
M.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24479  
A:Accession: T47902  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-248 <CHO>  
A:Cross-references: UNIPROT:Q9LZX3; UNIPARC:UPI000009C7A5; EMBL:AL162295  
A:Experimental source: cultivar Columbia; BAC clone T4C21  
C:Genetics:  
A:Map position: 3  
A:Introns: 61/3; 96/1; 119/1; 125/3; 143/2; 153/3; 181/1  
A:Note: T4C21.320

Query Match 70.8%; Score 17; DB 2; Length 248;  
Best Local Similarity 25.0%; Pred. No. 1.8e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 191 CSSAKSC 198

RESULT 27  
B87269  
exopolysaccharide production protein pss [imported] - *Caulobacter crescentus*  
C:Species: *Caulobacter crescentus*  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 31-Dec-2004  
C:Accession: B87269  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kjolton, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.; Proc. Natl. Acad. Sci. U.S.A. 98, 4135-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: E87269

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-267 <SNO>

A:Cross-references: UNIPROT:Q9ABR0; UNIPARC:UPI000000C6F59; GB:AE005673; NID:g13421281; F

C:Genetics:

A:Gene: CC0166

C:Superfamily: sugar transferase, ExoY type

Query Match 70.8%; Score 17; DB 2; Length 267;

Best Local Similarity 25.0%; Pred. No. 1.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8

DB 20 CTTSSDC 27

RESULT 28

F70771

probable glutamateracemase - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004

C:Accession: F70771

R:Coile, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sultson, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: F70771

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-271 <COL>

A:Cross-references: UNIPROT:Q10626; UNIPARC:UPI000012FA0E; GB:Z73902; GB:AL123456; NID:9

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: murI

C:Superfamily: glutamate racemase

Query Match 70.8%; Score 17; DB 2; Length 271;

Best Local Similarity 25.0%; Pred. No. 1.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8

DB 75 CNSASAC 82

RESULT 29

S72790

glutamate racemase (EC 5.1.1.3) - Mycobacterium leprae

N:Alternate names: B1549\_C2\_210 protein

C:Species: Mycobacterium leprae

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C:Accession: S72790

R:Smith, D.R.; Robison, K. submitted to the EMBL Data Library, November 1993

A:Description: Mycobacterium leprae cosmid B1549.

A:Reference number: S72582

A:Accession: S72790

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-272 <SML>

A:Cross-references: UNIPROT:P46705; UNIPARC:UPI000012FA0D; EMBL:U00014; NID:g466903; PID

C:Genetics:

A:Gene: murI

C:Superfamily: glutamate racemase

C:Keywords: isomerase

Query Match 70.8%; Score 17; DB 1; Length 272;

Best Local Similarity 25.0%; Pred. No. 1.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8

DB 75 CNTASAC 82

RESULT 30

I51172

transcription factor Rcc/EBP-2 - bullfrog

C:Species: Rana catesbeiana (bullfrog)

C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004

C:Accession: I51172

R:Chen, Y.; Hu, H.; Atkinson, B.G. Dev. Genet. 15, 366-377, 1994

A:Title: Characterization and expression of C/EBP-like genes in the liver of *Rana catesbe*

A:Reference number: I51171; MUID:95008407; PMID:7923939

A:Accession: I51172

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-284 <CHE>

A:Cross-references: UNIPROT:Q91295; UNIPARC:UPI00000FACAC; EMBL:U08605; NID:g478890; PID

C:Superfamily: CCAAT/enhancer-binding protein alpha

Query Match 70.8%; Score 17; DB 2; Length 284;

Best Local Similarity 25.0%; Pred. No. 1.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8

DB 171 CSMTSAC 178

RESULT 31

C64795

citG protein - *Escherichia coli* (strain K-12)

C:Species: *Escherichia coli*

C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004

C:Accession: C64795

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A:Rose, D.J.; Mau, B.; Sano, Y. Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: C64795

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-292 <BLAT>

A:Cross-references: UNIPROT:P77231; UNIPARC:UPI0000127A40; GB:AE000166; GB:U00096; NID:91

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: citG

C:Superfamily: citG protein

Query Match 70.8%; Score 17; DB 2; Length 292;

Best Local Similarity 25.0%; Pred. No. 1.9e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8

DB 142 CSTASFC 149

RESULT 32

D90710

2-(5'-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase (EC 4.2.-.-) *Escherich*

C:Species: *Escherichia coli*

C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004

C:Accession: D90710  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
 Gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shingana, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc  
 A:Reference number: A96629; MUID:21156231; PMID:11258796  
 A:Accession: D90710  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-292 <HAY>  
 A:Cross-references: UNIPROT:P58161; UNIPARC:UP10000127A3F; GB:BA000007; PIDN:BA034075.1;  
 A:Experimental source: strain O157:H7, substrain RMD 0509952  
 C:Genetics:  
 A:Gene: BC0652  
 C:Superfamily: citG protein  
 C:Keywords: carbon-oxygen lyase

Query Match      70.8%    Score 17; DB 2; Length 292;  
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
 DB 142 CSTASFC 149

RESULT 33

H8560  
 2-(5'-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase (EC 4.2.-.-) citG - *Escherichia*  
 C:Species: *Escherichia coli*  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: H8560  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Diallanita, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: H8560  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-292 <STO>  
 A:Cross-references: UNIPROT:P58161; UNIPARC:UP10000127A3F; GB:AE005174; NID:G12513512; F  
 A:Experimental source: strain O157:H7, substrain EDJ933  
 C:Genetics:  
 A:Gene: citG  
 C:Superfamily: citG protein  
 C:Keywords: carbon-oxygen lyase

Query Match      70.8%    Score 17; DB 2; Length 292;  
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
 DB 142 CSTASFC 149

RESULT 34

A98294  
 hypothetical protein AGR\_L\_2619 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere  
 C:Species: *Agrobacterium tumefaciens*  
 C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
 C:Accession: A98294  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*  
 A:Reference number: A97359; MUID:21608551; PMID:11743194  
 A:Accession: A98294  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-297 <KDU>  
 A:Cross-references: UNIPROT:Q8UA53; UNIPARC:UP100000D222C; GB:AE007870; PIDN:AAK89875.1;

C:Genetics:  
 A:Gene: AGR\_L\_2619  
 A:Map position: linear chromosome

Query Match      70.8%    Score 17; DB 2; Length 297;  
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
 DB 151 CAAVAAC 158

RESULT 35

AH2989  
 permease [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)  
 C:Species: *Agrobacterium tumefaciens*  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
 C:Accession: AH2989  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; Mciclell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
 A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: AH2989  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-297 <KDU>  
 A:Cross-references: UNIPROT:Q8UA53; UNIPARC:UP100000D222C; GB:AE008689; PIDN:AAL44334.1;  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Adu3521  
 A:Map position: linear chromosome

Query Match      70.8%    Score 17; DB 2; Length 297;  
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
 DB 151 CAAVAAC 158

RESULT 36

AE0857  
 conserved hypothetical protein STY3065 [imported] - *Salmonella enterica* subsp. *enterica*  
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh  
 A>Note: this species has also been called *Salmonella typhi*  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C:Accession: AE0857  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
 ; S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov.  
 A:Reference number: AB0502; MUID:21534947; PMID:11677608  
 A:Accession: AE0857  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-305 <PAR>  
 A:Cross-references: UNIPARC:UP1000005A39C; GB:AL513382; PIDN:CAD06044.1; PID:G16504011;  
 C:Genetics:  
 A:Gene: STY3065  
 C:Superfamily: *Escherichia coli* hypothetical protein b2755

Query Match      70.8%    Score 17; DB 2; Length 305;  
 Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
Db 179 CISAATSC 186

RESULT 37  
hypotheoretical protein b2755 - Escherichia coli (strain K-12)  
G65056  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
C:Accession: G65056  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; PMID:97426617; PMID:9278503  
A:Accession: G65056  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-305 <BLAT>  
A:Cross-references: UNIPROT:Q46896; UNIPARC:UPI000013B034; GB:AE000359; GB:U00096; NID:G  
C:Experimental source: strain K-12, substrain MG1655  
C:Superfamily: Escherichia coli hypotheoretical protein b2755

Query Match 70.8%; Score 17; DB 2; Length 305;  
Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
Db 179 CISAATSC 186

RESULT 38  
hypotheoretical protein Z4064 [imported] - Escherichia coli (strain O157:H7, substrain EDL9  
B85925  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: B85925  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glanier, J.D.; Rose, D.J.; Mayhew  
11ler, L.; Grobeck, B.J.; Davis, N.W.; Lim, A.; Diallanca, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; PMID:21074935; PMID:11206551  
A:Accession: B85925  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-307 <STO>  
A:Cross-references: UNIPROT:Q8X7W3; UNIPARC:UPI00000D09E4; GB:AE005174; NID:g12517212; F  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z4064  
C:Superfamily: Escherichia coli hypotheoretical protein b2755

Query Match 70.8%; Score 17; DB 2; Length 307;  
Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
Db 179 CISAATSC 186

RESULT 39  
A91080  
hypotheoretical protein ECs3609 [imported] - Escherichia coli (strain O157:H7, substrain RI  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: A91080  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend

A:Reference number: A99629; PMID:21156231; PMID:11258796  
A:Accession: A91080  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-307 <HAV>  
A:Cross-references: UNIPROT:Q8X7W3; UNIPARC:UPI00000D09E4; GB:BA000007; PIDN:BA837032.1;  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECs3609  
C:Superfamily: Escherichia coli hypotheoretical protein b2755

Query Match 70.8%; Score 17; DB 2; Length 307;  
Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
Db 179 CISAATSC 186

RESULT 40  
T32376  
hypotheoretical protein K10F12.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T32376  
R:Wohlmann, P.; Beck, C.  
submitted to the EMBL Data Library, September 1997  
A:Description: The sequence of C. elegans cosmid K10F12.  
A:Reference number: Z21157  
A:Accession: T32376  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-309 <NOH>  
A:Cross-references: UNIPROT:O17234; UNIPARC:UPI000007F612; EMBL:AF025462; PIDN:AAB71002.1  
A:Experimental source: strain Bristol N2; clone K10F12  
C:Genetics:  
A:Gene: CESP:K10F12.4  
A:Map position: 3  
A:Insertions: 31/3; 123/2; 196/3; 239/1

Query Match 70.8%; Score 17; DB 2; Length 309;  
Best Local Similarity 25.0%; Pred. No. 1.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
Db 7 CTSALC 14

RESULT 41  
B53522  
20K cyclophilin - Toxoplasma gondii (fragment)  
C:Species: Toxoplasma gondii  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: B53522  
R:High, K.P.; Joiner, K.A.; Handschumacher, R.E.  
J Biol. Chem. 269, 9105-9112, 1994  
A:Title: Isolation, cDNA sequences, and biochemical characterization of the major cyclophi  
A:Reference number: A53522; PMID:94179329; PMID:8132648  
A:Accession: B53522  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-347 <HTG>  
A:Cross-references: UNIPROT:Q26995; UNIPARC:UPI000007DB6A; GB:U04634; NID:g436958; PID:g'  
F:178-346/Domain: cyclophilin homology <CTP>

Query Match 70.8%; Score 17; DB 2; Length 347;  
Best Local Similarity 25.0%; Pred. No. 2e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8



Db 48 CASAEAC 55

## RESULT 42

probable iron-siderophore uptake system transmembrane component - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C/Accession: T36890

R/Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A/Reference number: 221617

A/Accession: T36890

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-348 <MUR>

A/Cross-references: UNIPROT:O9S213; UNIPARC:UPI00000DB327; EMBL:AL109848; PIDN:CA52851.

A/Experimental source: strain A3(2)

C/Genetics:

C/Superfamily: vitamin B12 transport protein btuc

Query Match 70.8%; Score 17; DB 2; Length 348;  
Best Local Similarity 25.0%; Pred. No. 2e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 262 CGAATAC 269

## RESULT 43

hypothetical protein At2g17120 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C/Accession: C84548

R/Li, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.

et al.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: C84548

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-350 <STO>

A/Cross-references: UNIPROT:O23006; UNIPARC:UPI00000A22DD; GB:AE002093; NID:g2558660; PI

C/Genetics:

A/Gene: At2g17120

A/Map position: 2

Query Match 70.8%; Score 17; DB 2; Length 350;  
Best Local Similarity 25.0%; Pred. No. 2e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 31 CGSGTSTC 38

## RESULT 44

conserved hypothetical protein CC2083 [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus

C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004

C/Accession: B87507

R/Nierman, W.C.; Feldlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolor

n, U.; Emolvaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of Caulobacter crescentus.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: B87507  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-352 <STO>  
A/Cross-references: UNIPROT:Q9A6L0; UNIPARC:UPI00000C75D3; GB:AE005673; NID:g13423564; P  
C/Genetics:  
A/Gene: CC2083

Query Match 70.8%; Score 17; DB 2; Length 352;  
Best Local Similarity 25.0%; Pred. No. 2e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 11 CAALISAC 18

## RESULT 45

conserved hypothetical protein TP0730 - syphilis spirochete

C/Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C/Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004

C/Accession: G71287

R/Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin

et al.; Kralak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uitterlinden, T.; McDo

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A/Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A/Reference number: A71250; MUID:98332770; PMID:9665876

A/Accession: G71287

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-365 <COL>

A/Cross-references: UNIPROT:O83712; UNIPARC:UPI00000C0AA2; GB:AE001245; GB:AE000520; NID

A/Experimental source: strain Nichols

C/Genetics:

A/Gene: TP0730

Query Match 70.8%; Score 17; DB 2; Length 365;  
Best Local Similarity 25.0%; Pred. No. 2e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
Db 162 CAASDC 169

## RESULT 46

nifs family enzyme (cysteine desulfurase/cysteine sulfinate desulfinase) [imported] - Cl

C/Species: Clostridium acetobutylicum

C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004

C/Accession: D97175

R/Nolling, U.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J.; Daily, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A/Reference number: A96900; MUID:21359325; PMID:21359325

A/Accession: D97175

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-376 <KUR>

A/Cross-references: UNIPROT:O97GY1; UNIPARC:UPI00000D7546; GB:AE001437; PIDN:AAK80191.1;

A/Experimental source: Clostridium acetobutylicum ATCC824

C/Genetics:

A/Gene: CAC2234

C/Superfamily: nitrogen fixation protein nifs

Query Match 70.8%; Score 17; DB 2; Length 376;  
Best Local Similarity 25.0%; Pred. No. 2e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8

Db 319 CASGSAC 326

## RESULT 47

C34443

nitrogenase cofactor synthesis protein nifs - Anabaena sp.

N:Contains: L-cysteine sulfotransferase (EC 2.8.1.-)

C:Species: Anabaena sp.

C&gt;Date: 15-Jun-1990 #sequence\_revision 15-Jun-1990 #text\_change 09-Jul-2004

C/Accession: C34443; B32361

R:Muligan, M.E.; Haselkorn, R.

J. Biol. Chem. 264, 19200-19207, 1989

A&gt;Title: Nitrogen fixation (nif) genes of the cyanobacterium Anabaena species strain PCC

A/Reference number: A34443; MUID:90037054; PMID:2553733

A/Accession: C34443

A&gt;Status: preliminary

A/Molecule type: DNA

A/Residues: 1-400 &lt;MUL&gt;

A/Cross-references: UNIPROT:P12623; UNIPARC:UPI000016555C; GB:J05111; NID:9142034; PIDN:

R:Muligan, M.E.; Bullock, W.J.; Haselkorn, R.

J. Bacteriol. 170, 4406-4410, 1988

A&gt;Title: Bacterial-type ferredoxin genes in the nitrogen fixation regions of the cyanoba

A/Reference number: A31884; MUID:88314954; PMID:2842320

A/Accession: B32361

A/Molecule type: DNA

A/Residues: 1-7 &lt;MUL&gt;

A/Cross-references: UNIPARC:UPI0000178A6F

C/Superfamily: nitrogen fixation protein nifs

C/Keywords: phosphoprotein; pyridoxal phosphate; sulfotransferase

F/324/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

F/324/Active site: Cys (cysteine persulfide intermediate) #status predicted

Query Match 70.8%; Score 17; DB 2; Length 400;

Best Local Similarity 25.0%; Pred. No. 2e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8

Db 317 CASGSAC 324

## RESULT 48

AE188

nitrogenase cofactor synthesis protein (imported) - Nostoc sp. (strain PCC 7120)

C/Species: Nostoc sp. PCC 7120

A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004

C/Accession: AE1988

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriguchi,

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A&gt;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A/Reference number: AB1807; MUID:21595285; PMID:11759840

A/Accession: AE1988

A&gt;Status: preliminary

A/Molecule type: DNA

A/Residues: 1-400 &lt;KUR&gt;

A/Cross-references: UNIPROT:P12623; UNIPARC:UPI000013010C; GB:BA000019; PIDN:BA073413.1;

A/Experimental source: strain PCC 7120

C/Genetics:

A/Gene: nifs

C/Superfamily: nitrogen fixation protein nifs

Query Match 70.8%; Score 17; DB 2; Length 400;

Best Local Similarity 25.0%; Pred. No. 2e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8

Db 317 CASGSAC 324

## RESULT 49

WNVIA

immediate-early protein IE-N - Autographa californica nuclear polyhedrosis virus

C/Species: Autographa californica nuclear polyhedrosis virus, AcNPV

C/Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004

C/Accession: A39150

R:Carson, D.D.; Summers, M.D.; Guarino, L.A.

Virology 182, 279-286, 1991

A&gt;Title: Molecular analysis of a baculovirus regulatory gene.

A/Reference number: A39150; MUID:91220660; PMID:2024466

A/Accession: A39150

A/Molecule type: DNA

A/Residues: 1-408 &lt;CAR&gt;

A/Cross-references: UNIPROT:P24647; UNIPARC:UPI0000170DEB; GB:M59422; NID:9332437; PIDN:

C/Superfamily: AcNPV immediate-early protein IE-N

C/Keywords: DNA binding; early protein; tandem repeat; transcription regulation

F/34-49/Region: 7-residue repeats

F/51-58/Region: 4-residue repeats

F/190-196/Region: glutamine-rich

Query Match 70.8%; Score 17; DB 1; Length 408;

Best Local Similarity 25.0%; Pred. No. 2.1e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8

Db 253 CSATSDC 260

## RESULT 50

A72869

early gene transactivator - Autographa californica nuclear polyhedrosis virus

C/Species: Autographa californica nuclear polyhedrosis virus, AcNPV

A/Note: dsDNA virus

C/Date: 12-Nov-1999 #sequence\_revision 12-Nov-1999 #text\_change 09-Jul-2004

C/Accession: A72869

R:Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.

Virology 202, 586-605, 1994

A&gt;Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.

A/Reference number: A72850; MUID:94303173; PMID:8030224

A/Accession: A72869

A&gt;Status: preliminary

A/Molecule type: DNA

A/Residues: 1-408 &lt;AYR&gt;

A/Cross-references: UNIPROT:P24647; UNIPARC:UPI00001387E9; GB:L22858; NID:9510708; PIDN:

C/Genetics:

C/Superfamily: AcNPV immediate-early protein IE-N

Query Match 70.8%; Score 17; DB 2; Length 408;

Best Local Similarity 25.0%; Pred. No. 2.1e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8

Db 253 CSATSDC 260

Search completed: January 4, 2006, 16:10:26  
Job time : 9.27826 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 / Search time 52.2435 seconds  
(without alignments)  
108.037 Million cell updates/sec

Title: US-09-932-322-9  
Perfect score: 24  
Sequence: 1 CXXXXXXC 8

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database: UniProt\_05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	75.0	69	2 Q9N669_9CAEN	Q9N669 conus ebrae
2	18	75.0	69	2 Q9NCS8_9CAEN	Q9NCS8 conus ebrae
3	18	75.0	69	2 Q9NCS9_9CAEN	Q9NCS9 conus ebrae
4	18	75.0	69	2 Q9NCT2_9CAEN	Q9NCT2 conus ebrae
5	18	75.0	69	2 Q9NCT3_9CAEN	Q9NCT3 conus ebrae
6	18	75.0	69	2 Q9VD36_DROME	Q9VD36 dirosophila
7	18	75.0	90	2 Q9FPP9_ANGOA	Q9FPP9 anopheles g
8	18	75.0	90	2 Q9B111_ANGOA	Q9B111 anopheles g
9	18	75.0	114	2 Q82854_9RETR	Q82854 jembrana di
10	18	75.0	161	2 Q9EQJ8_MOUSE	Q9EQJ8 mus musculu
11	18	75.0	163	2 Q7QPM4_GIALA	Q7QPM4 giardia lam
12	18	75.0	177	2 Q8LDT9_ARATH	Q8LDT9 arabidopsis
13	18	75.0	177	2 Q9C9L7_ARATH	Q9C9L7 arabidopsis
14	18	75.0	199	2 Q6OM47_CAEBR	Q6OM47 caenorhabdi
15	18	75.0	200	2 Q5VMP6_ORYSA	Q5VMP6 oryza sativ
16	18	75.0	224	2 Q9EQJ7_MOUSE	Q9EQJ7 mus musculu
17	18	75.0	233	2 Q6K2G5_ORYSA	Q6K2G5 oryza sativ
18	18	75.0	242	2 Q51XD3_MAGGR	Q51XD3 magnaporthe
19	18	75.0	245	2 Q6BCP3_HPBVO	Q6BCP3 hepaticis b
20	18	75.0	258	2 Q9YOE9_DROME	Q9YOE9 dirosophila
21	18	75.0	269	2 Q7PRQ7_ANGOA	Q7PRQ7 anopheles g
22	18	75.0	319	2 Q7Y198_ORYSA	Q7Y198 oryza sativ
23	18	75.0	320	2 Q61AJ5_CAEBR	Q61AJ5 caenorhabdi
24	18	75.0	328	2 Q6ZNM3_HUMAN	Q6ZNM3 homo sapien
25	18	75.0	332	2 Q6Z4R6_ORYSA	Q6Z4R6 oryza sativ
26	18	75.0	350	2 Q94589_9CTIL	Q94589 lambdaion b
27	18	75.0	355	1 RHOM_DROME	RHOM dirosophila
28	18	75.0	355	1 Q540V7_DROME	Q540V7 dirosophila
29	18	75.0	367	1 T1SD_MOUSE	T1SD mus musculu
30	18	75.0	404	2 Q665B6_DROVI	Q665B6 dirosophila
31	18	75.0	417	2 Q6Z4R5_ORYSA	Q6Z4R5 oryza sativ

32	18	75.0	445	2 Q8UTL4_ORYLA	Q8UTL4 oryzias lat
33	18	75.0	480	2 Q7SHC6_NEUCR	Q7SHC6 neurospora
34	18	75.0	480	2 Q758Y3_ASHCO	Q758Y3 ashbya goss
35	18	75.0	492	1 T1SD_HUMAN	T1SD homo sapien
36	18	75.0	494	2 Q53TE4_HUMAN	Q53TE4 homo sapien
37	18	75.0	526	2 Q6PD21_MOUSE	Q6PD21 mus musculu
38	18	75.0	544	2 Q5B195_DROME	Q5B195 dirosophila
39	18	75.0	545	2 Q91TR2_TUHV1	Q91TR2 tupaiid her
40	18	75.0	551	2 Q5Z8N9_ORYSA	Q5Z8N9 oryza sativ
41	18	75.0	551	2 Q8SB30_ORYSA	Q8SB30 oryza sativ
42	18	75.0	552	2 Q5AP87_CANAL	Q5AP87 candida alb
43	18	75.0	560	2 Q8UJ53_HPBVO	Q8UJ53 hepaticis b
44	18	75.0	581	2 Q6N1S1_CORDI	Q6N1S1 corneabacte
45	18	75.0	592	2 Q51VZ0_MAGGR	Q51VZ0 magnaporthe
46	18	75.0	606	2 Q6Z4E7_ORYSA	Q6Z4E7 oryza sativ
47	18	75.0	691	2 Q4RNV9_TETNG	Q4RNV9 tetradon n
48	18	75.0	705	2 Q7S136_NEUCR	Q7S136 neurospora
49	18	75.0	711	2 Q5RPM2_ERARE	Q5RPM2 brachydanio
50	18	75.0	722	2 Q5AP84_CANAL	Q5AP84 candida alb
51	18	75.0	743	2 Q09517_HPBVO	Q09517 hepaticis b
52	18	75.0	769	2 Q8QGX5_HPBVO	Q8QGX5 hepaticis b
53	18	75.0	787	2 Q09511_HPBVO	Q09511 hepaticis b
54	18	75.0	795	2 Q8V1H4_HPBVO	Q8V1H4 hepaticis b
55	18	75.0	799	2 Q4T7H8_TETNG	Q4T7H8 tetradon n
56	18	75.0	801	2 Q09504_HPBVO	Q09504 hepaticis b
57	18	75.0	801	2 Q09505_HPBVO	Q09505 hepaticis b
58	18	75.0	823	2 Q9QMN3_HPBVO	Q9QMN3 hepaticis b
59	18	75.0	835	2 Q8V1I3_HPBVO	Q8V1I3 hepaticis b
60	18	75.0	837	2 Q5OUT2_HPBVO	Q5OUT2 hepaticis b
61	18	75.0	843	1 DPOL_HPBVO	DPOL_HPBVO
62	18	75.0	843	1 Q09509_HPBVO	Q09509 hepaticis b
63	18	75.0	843	2 Q39877_HPBVO	Q39877 hepaticis b
64	18	75.0	843	2 Q39882_HPBVO	Q39882 hepaticis b
65	18	75.0	843	2 Q91527_HPBVO	Q91527 hepaticis b
66	18	75.0	843	2 Q5DW06_HPBVO	Q5DW06 hepaticis b
67	18	75.0	843	2 Q5DW09_HPBVO	Q5DW09 hepaticis b
68	18	75.0	843	2 Q5DM13_HPBVO	Q5DM13 hepaticis b
69	18	75.0	843	2 Q5Q0T5_HPBVO	Q5Q0T5 hepaticis b
70	18	75.0	843	2 Q67937_HPBVO	Q67937 hepaticis b
71	18	75.0	843	2 Q68RRT_HPBVO	Q68RRT hepaticis b
72	18	75.0	843	2 Q68RP0_HPBVO	Q68RP0 hepaticis b
73	18	75.0	843	2 Q68RP3_HPBVO	Q68RP3 hepaticis b
74	18	75.0	843	2 Q68RQ6_HPBVO	Q68RQ6 hepaticis b
75	18	75.0	843	2 Q68RQ9_HPBVO	Q68RQ9 hepaticis b
76	18	75.0	843	2 Q68RQ3_HPBVO	Q68RQ3 hepaticis b
77	18	75.0	843	2 Q762E4_HPBVO	Q762E4 hepaticis b
78	18	75.0	843	2 Q765V6_HPBVO	Q765V6 hepaticis b
79	18	75.0	843	2 Q76B37_HPBVO	Q76B37 hepaticis b
80	18	75.0	843	2 Q77DP5_HPBVO	Q77DP5 hepaticis b
81	18	75.0	843	2 Q77DP8_HPBVO	Q77DP8 hepaticis b
82	18	75.0	843	2 Q77DP8_HPBVO	Q77DP8 hepaticis b
83	18	75.0	843	2 Q77DQ1_HPBVO	Q77DQ1 hepaticis b
84	18	75.0	843	2 Q77DQ7_HPBVO	Q77DQ7 hepaticis b
85	18	75.0	843	2 Q77DR0_HPBVO	Q77DR0 hepaticis b
86	18	75.0	843	2 Q77DS1_HPBVO	Q77DS1 hepaticis b
87	18	75.0	843	2 Q77DP5_HPBVO	Q77DP5 hepaticis b
88	18	75.0	843	2 Q77DP5_HPBVO	Q77DP5 hepaticis b
89	18	75.0	843	2 Q77H00_HPBVO	Q77H00 hepaticis b
90	18	75.0	843	2 Q77H00_HPBVO	Q77H00 hepaticis b
91	18	75.0	843	2 Q77H00_HPBVO	Q77H00 hepaticis b
92	18	75.0	843	2 Q805G6_HPBVO	Q805G6 hepaticis b
93	18	75.0	843	2 Q80GVR_HPBVO	Q80GVR hepaticis b
94	18	75.0	843	2 Q80H00_HPBVO	Q80H00 hepaticis b
95	18	75.0	843	2 Q80H20_HPBVO	Q80H20 hepaticis b
96	18	75.0	843	2 Q80H48_HPBVO	Q80H48 hepaticis b
97	18	75.0	843	2 Q80U60_HPBVO	Q80U60 hepaticis b
98	18	75.0	843	2 Q80J78_HPBVO	Q80J78 hepaticis b
99	18	75.0	843	2 Q80J78_HPBVO	Q80J78 hepaticis b
100	18	75.0	843	2 Q80J83_HPBVO	Q80J83 hepaticis b
101	18	75.0	843	2 Q80M08_HPBVO	Q80M08 hepaticis b
102	18	75.0	843	2 Q81116_HPBVO	Q81116 hepaticis b
103	18	75.0	843	2 Q81120_HPBVO	Q81120 hepaticis b
104	18	75.0	843	2 Q81120_HPBVO	Q81120 hepaticis b

105	18	75.0	843	2	Q81124_HPBVO	Q81124 hepatitis b	178	17	70.8	120	2	Q49859_MYCLE	Q49859 mycobacteri
106	18	75.0	843	2	Q8A240_HPBVO	Q8a240 hepatitis b	179	17	70.8	123	2	Q8A174_CAEBL	Q8a174 caenorhabdi
107	18	75.0	843	2	Q8OSD1_HPBVO	Q8osd1 hepatitis b	180	17	70.8	124	2	Q56R10_PENMO	Q56r10 penaeus mon
108	18	75.0	843	2	Q8VOM9_HPBVO	Q8vom9 hepatitis b	181	17	70.8	125	2	Q9B114_GIALA	Q9b114 giardia lam
109	18	75.0	843	2	Q8VONI_HPBVO	Q8von1 hepatitis b	182	17	70.8	126	2	Q9B119_GIALA	Q9b119 giardia lam
110	18	75.0	843	2	Q9IF27_HPBVO	Q9if27 hepatitis b	183	17	70.8	126	2	Q9B1M3_GIALA	Q9b1m3 giardia lam
111	18	75.0	843	2	Q9QAB8_HPBVO	Q9qab8 hepatitis b	184	17	70.8	127	2	Q5TOD0_ANOGA	Q5tdo0 anopheles g
112	18	75.0	843	2	Q9QAC4_HPBVO	Q9qac4 hepatitis b	185	17	70.8	127	2	Q9B1K7_GIALA	Q9b1k7 giardia lam
113	18	75.0	843	2	Q9QAD4_HPBVO	Q9qad4 hepatitis b	186	17	70.8	128	2	Q9B1K6_GIALA	Q9b1k6 giardia lam
114	18	75.0	843	2	Q9QAE0_HPBVO	Q9qae0 hepatitis b	187	17	70.8	128	2	Q9B1K9_GIALA	Q9b1k9 giardia lam
115	18	75.0	843	2	Q9QBE2_HPBVO	Q9qbe2 hepatitis b	188	17	70.8	128	2	Q74A10_GEOSL	Q74a10 geobacter b
116	18	75.0	843	2	Q9QBE7_HPBVO	Q9qbe7 hepatitis b	189	17	70.8	129	2	Q5VJF5_RAT	Q5vjf5 ratulus norv
117	18	75.0	843	2	Q9QBF1_HPBVO	Q9qbf1 hepatitis b	190	17	70.8	130	1	Q5PHB5_HUMAN	Q5phb5 homo sapien
118	18	75.0	843	2	Q9QBE6_HPBVO	Q9qbe6 hepatitis b	191	17	70.8	130	1	Q5PHB5_MOUSE	Q5phb5 mus musculu
119	18	75.0	843	2	Q9Y194_HPBVO	Q9y194 hepatitis b	192	17	70.8	130	2	Q6NTD0_HUMAN	Q6ntd0 homo sapien
120	18	75.0	843	2	Q6R264_HPBVO	Q6r264 hepatitis b	193	17	70.8	130	2	Q6AVS9_ORYSA	Q6avs9 oryza sativ
121	18	75.0	843	2	Q7THS7_HPBVO	Q7ths7 hepatitis b	194	17	70.8	131	2	Q5T0Y4_HUMAN	Q5t0y4 homo sapien
122	18	75.0	843	2	Q8UJ87_HPBVO	Q8uj87 hepatitis b	195	17	70.8	132	2	Q8PKN7_XANNC	Q8pkn7 xanthomonas
123	18	75.0	843	2	Q5UAY3_HPBVO	Q5uay3 hepatitis b	196	17	70.8	135	1	TVBI_HUMAN	TVBI_HUMAN
124	18	75.0	843	2	Q67925_HPBVO	Q67925 hepatitis b	197	17	70.8	139	2	Q58508_PYRHO	Q58508 pyrococcus
125	18	75.0	843	2	Q68RS2_HPBVO	Q68rs2 hepatitis b	198	17	70.8	140	2	Q6U7V1_CRIBE	Q6u7v1 crinipellis
126	18	75.0	843	2	Q6JWV7_HPBVO	Q6jwv7 hepatitis b	199	17	70.8	140	2	Q58HK1_HUMAN	Q58hk1 homo sapien
127	18	75.0	843	2	Q9PK62_HPBVO	Q9pk62 hepatitis b	200	17	70.8	140	2	Q68DC0_HUMAN	Q68dc0 homo sapien
128	18	75.0	843	2	Q9QAD0_HPBVO	Q9qad0 hepatitis b	201	17	70.8	140	2	Q6F369_ORYSA	Q6f369 oryza sativ
129	18	75.0	843	2	Q9QAB8_HPBVO	Q9qab8 hepatitis b	202	17	70.8	145	2	Q8VWK7_GIAMI	Q8vwk7 antirrhinum
130	18	75.0	843	2	Q9QMM9_HPBVO	Q9qmm9 hepatitis b	203	17	70.8	145	2	Q8W019_GIAMI	Q8w019 misopates o
131	18	75.0	843	2	Q5UJUS_HPBVO	Q5ujus hepatitis b	204	17	70.8	145	2	Q8W020_GIAMI	Q8w020 misopates o
132	18	75.0	845	2	Q4KRS9_HPBVO	Q4krs9 hepatitis b	205	17	70.8	145	2	Q8W021_GIAMI	Q8w021 antirrhinum
133	18	75.0	845	2	Q52EW8_MAGCR	Q52ew8 magnaporthe	206	17	70.8	145	2	Q8W022_ANTMA	Q8w022 antirrhinum
134	18	75.0	984	2	Q4SKG0_TETNG	Q4skg0 tetradodon n	207	17	70.8	145	2	Q8W023_ANTMA	Q8w023 antirrhinum
135	18	75.0	1135	2	Q4PBA7_USTMA	Q4pba7 ubtillago ma	208	17	70.8	145	2	Q9AB81_CAOCR	Q9ab81 caulobacter
136	18	75.0	1262	2	Q4RLJ3_TETNG	Q4rlj3 tetradodon n	209	17	70.8	146	2	Q9DDE1_BRARE	Q9dde1 brachydanio
137	18	75.0	1286	2	Q9VRA9_DROME	Q9vra9 drosophila	210	17	70.8	148	2	Q4UPP4_XANCP	Q4upp4 xanthomonas
138	18	75.0	1348	2	Q4QJDO_LEIMA	Q4qjdo leishmania	211	17	70.8	148	2	Q8P460_XANCP	Q8p460 xanthomonas
139	18	75.0	1349	2	Q675Y9_QUIROC	Q675y9 okpopleura	212	17	70.8	150	2	Q7PTR6_ANOGA	Q7ptr6 anopheles g
140	18	75.0	1510	2	Q8PKK6_XANNC	Q8pkk6 xanthomonas	213	17	70.8	153	2	Q6MZK7_HUMAN	Q6mzk7 homo sapien
141	18	75.0	1743	2	Q9XWXS_CAEBL	Q9xwxs caenorhabdi	214	17	70.8	156	2	Q69S06_ORYSA	Q69s06 oryza sativ
142	18	75.0	2543	2	P90649_PAPRP	P90649 plasmecium	215	17	70.8	157	2	Q4LXB6_9BURK	Q4lxb6 burkholderi
143	18	75.0	4261	2	Q8IFP4_PLAF7	Q8ifp4 piastecium	216	17	70.8	158	2	Q9B1J9_GIALA	Q9b1j9 giardia lam
144	18	70.8	26	2	Q93940_PODAN	Q93940 podopora a	217	17	70.8	158	2	Q6Z4C9_ORYSA	Q6z4c9 oryza sativ
145	17	70.8	26	2	Q9TRM6_BOVIN	Q9trm6 bos taurus	218	17	70.8	161	2	Q9B1J8_GIALA	Q9b1j8 giardia lam
146	17	70.8	37	2	Q7UE39_RHOBA	Q7ue39 rhodopirell	219	17	70.8	161	2	Q4RZB3_TETNG	Q4rzb3 tetradodon n
147	17	70.8	49	2	Q4WHU3_NSPFU	Q4whu3 aspergillus	220	17	70.8	162	2	Q91GT6_ADEP3	Q91gt6 porcine ade
148	17	70.8	59	2	Q7S2A3_NEUCR	Q7s2a3 neurospora	221	17	70.8	164	2	Q77790_BOVIN	Q77790 bos taurus
149	17	70.8	62	2	Q84TW7_ORYSA	Q84tw7 oryza sativ	222	17	70.8	166	2	Q4M038_9BURK	Q4m038 burkholderi
150	17	70.8	67	2	Q7QWK1_GIALA	Q7qwk1 giardia lam	223	17	70.8	168	2	Q9ZKFS_RHIME	Q9zkfs rhizobium m
151	17	70.8	83	2	Q7R273_GIALA	Q7r273 giardia lam	224	17	70.8	174	1	ITRY_BATVA	ITRY_BATVA
152	17	70.8	84	2	Q8GWP9_ARATH	Q8gwp9 arabidopsis	225	17	70.8	175	2	Q9SLZ1_SHEEP	Q9slz1 ovis aries
153	17	70.8	84	2	Q8L5N7_ARATH	Q8l5n7 arabidopsis	226	17	70.8	178	2	Q7NTE5_CHRVO	Q7nte5 chromobacte
154	17	70.8	85	2	Q5W46_LEGPL	Q5w46 legionella	227	17	70.8	179	2	Q61TY1_CAEBR	Q61ty1 caenorhabdi
155	17	70.8	85	2	Q5X4Q7_LEGPA	Q5x4q7 legionella	228	17	70.8	181	2	Q6Z6C6_ORYSA	Q6z6c6 oryza sativ
156	17	70.8	85	2	Q5ZUZ4_LEGPH	Q5zuz4 legionella	229	17	70.8	182	1	YAF2_CAEBL	YAF2_CAEBL
157	17	70.8	86	2	Q8DDM7_VIBVU	Q8ddm7 vibrio vuln	230	17	70.8	184	2	Q61ZFO_CAEBR	Q61zfo caenorhabdi
158	17	70.8	86	2	Q7MGN7_VIBVY	Q7mgn7 vibrio vuln	231	17	70.8	184	2	Q72SD2_DESVH	Q72sd2 desulfovibr
159	17	70.8	87	2	Q5UKM7_ORYSA	Q5ukm7 oryza sativ	232	17	70.8	185	2	Q8U0G0_TETNG	Q8u0g0 tetradodon n
160	17	70.8	88	2	Q74DJ1_GEOSL	Q74dj1 geobacter s	233	17	70.8	188	2	Q58G82_HUMAN	Q58g82 homo sapien
161	17	70.8	88	2	Q90ZP3_XENLA	Q90zpj xenopus lae	234	17	70.8	188	2	Q90XY9_SCYCA	Q90xy9 scyllorhino
162	17	70.8	89	2	Q49652_MYCLE	Q49652 mycobacteri	235	17	70.8	194	2	Q86MJB_MUSDO	Q86mj8 musca domes
163	17	70.8	95	2	Q7YTR5_CAEBL	Q7ytrs caenorhabdi	236	17	70.8	195	1	COMD3_MOUSE	COMD3_MOUSE
164	17	70.8	95	2	P94628_CITPR	P94628 citrobacter	237	17	70.8	195	2	Q5USW6_MOUSE	Q5usw6 mus musculu
165	17	70.8	96	2	Q7YTR6_CAEBL	Q7ytr6 caenorhabdi	238	17	70.8	196	2	Q5M988_XENLA	Q5m988 xenopus lae
166	17	70.8	96	2	Q4SDL9_TETNG	Q4sdl9 tetradodon n	239	17	70.8	196	2	Q4SMD6_TETNG	Q4smd6 tetradodon n
167	17	70.8	99	1	NICI_HUMAN	NICI_HUMAN	240	17	70.8	200	2	Q7R5C3_GIALA	Q7r5c3 giardia lam
168	17	70.8	99	2	Q6IAD7_HUMAN	Q6iad7 homo sapien	241	17	70.8	203	2	P95044_MYCTU	P95044 mycobacteri
169	17	70.8	99	2	Q5Y4U4_AGEOR	Q5y4u4 agelena ori	242	17	70.8	204	2	Q4LHS3_SBURK	Q4lhs3 burkholderi
170	17	70.8	101	2	Q48435_DPSPP	Q48435 bacteriopho	243	17	70.8	205	2	Q5TQW7_ANOGA	Q5tqw7 anopheles g
171	17	70.8	102	2	Q5A3U4_CANAL	Q5a3u4 candida alb	244	17	70.8	205	2	Q4RH93_TETNG	Q4rh93 tetradodon n
172	17	70.8	102	2	Q4S1V5_TETNG	Q4s1v5 tetradodon n	245	17	70.8	209	2	Q94828_TETTH	Q94828 tetrahymena
173	17	70.8	108	2	Q5DAP3_SCHUA	Q5dah3 schistosoma	246	17	70.8	210	2	Q4P148_USITMA	Q4p148 usutillago ma
174	17	70.8	108	2	Q5DCP1_SCHUA	Q5dcp1 schistosoma	247	17	70.8	210	2	Q4RV89_TETNG	Q4rv89 tetradodon n
175	17	70.8	108	2	Q75HC3_ORYSA	Q7shc3 oryza sativ	248	17	70.8	212	2	Q9RRS8_DEIRA	Q9rrs8 deinococcus
176	17	70.8	108	2	Q53CR5_GGAMA	Q53cr5 macaca fusc	249	17	70.8	213	2	Q8U0F5_TETNG	Q8u0f5 tetradodon n
177	17	70.8	109	2	Q7ULF7_MYCBO	Q7ulf7 mycobacteri	250	17	70.8	215	2	Q9HCA4_HUMAN	Q9hca4 homo sapien

251	17	70.8	216	2	Q4T5X5_TETNG	Q4E5X5 tetraodon n	324	17	70.8	292	1	CITG_ECO57	P56161 escherichia
252	17	70.8	217	2	Q8BV88_MOISE	Q8BV88 mus musculus	325	17	70.8	292	1	CITG_ECOL6	Q8F804 escherichia
253	17	70.8	218	2	Q95WY8_IXOSC	Q95WY8 ixodes scap	326	17	70.8	292	1	CITG_ECOL6	P77231 escherichia
254	17	70.8	219	2	Q4N1W0_THRPA	Q4N1W0 theileria p	327	17	70.8	292	2	Q7UDF6_SHIFL	Q7UDF6 shigella fl
255	17	70.8	220	2	Q4PMF5_IXOSC	Q4PMF5 ixodes scap	328	17	70.8	293	2	Q6N1K5_DROME	Q6N1K5 drosophila
256	17	70.8	220	2	Q4PMK0_IXOSC	Q4PMK0 ixodes scap	329	17	70.8	294	2	Q62ON2_CAEBR	Q62ON2 caenorhabdi
257	17	70.8	221	1	KR103_HUMAN	P60369 homo sapien	330	17	70.8	294	2	Q615G2_CAEBR	Q615G2 caenorhabdi
258	17	70.8	223	1	Q7WXL2_ALCEU	Q7WXL2 alcaligenes	331	17	70.8	294	2	Q4RPT5_TETNG	Q4RPT5 tetraodon n
259	17	70.8	224	1	MYOG_HUMAN	P15173 homo sapien	332	17	70.8	296	2	Q5PEG9_SALPA	Q5PEG9 salmonella
260	17	70.8	224	1	MYOG_PIG	P49812 sus scrofa	333	17	70.8	296	2	Q73TVO_MYCPA	Q73TVO mycobacteri
261	17	70.8	224	2	Q53XM6_HUMAN	Q53XM6 homo sapien	334	17	70.8	297	2	Q4P724_USTMA	Q4P724 ussiliago ma
262	17	70.8	224	2	Q5E4X4_CAEBL	Q5E4X4 caenorhabdi	335	17	70.8	297	2	Q4FX78_LEISHA	Q4FX78 leishmania
263	17	70.8	224	2	Q5EE03_PIG	Q5EE03 sus scrofa	336	17	70.8	297	2	Q8UA53_AGRIS	Q8UA53 agrobacteri
264	17	70.8	224	2	Q7YS81_BOVIN	Q7YS81 bos taurus	337	17	70.8	299	2	Q7Q3P3_ANOGA	Q7Q3P3 anopheles g
265	17	70.8	224	2	Q5EC23_PIG	Q5EC23 sus scrofa	338	17	70.8	299	2	Q4VSE5_DROME	Q4VSE5 drosophila
266	17	70.8	225	2	Q6BUR5_DEBHA	Q6BUR5 debrayomyce	339	17	70.8	300	2	Q4P6F7_USTMA	Q4P6F7 ussiliago ma
267	17	70.8	225	2	Q6YXK6_ORYSA	Q6YXK6 oryza sativ	340	17	70.8	301	2	Q5TOK9_ANOGA	Q5TOK9 anopheles g
268	17	70.8	226	2	Q5NYG7_AZOSE	Q5NYG7 azocarcus sp	341	17	70.8	301	2	Q8BN70_MOUSE	Q8BN70 mus musculus
269	17	70.8	227	2	Q5N970_ORYSA	Q5N970 oryza sativ	342	17	70.8	302	2	HXC1A_BRARE	Q98H9 brachydantio
270	17	70.8	227	2	Q51JQ3_9VIRU	Q51JQ3 banana viru	343	17	70.8	302	2	Q7XH78_ORYSA	Q7XH78 oryza sativ
271	17	70.8	230	1	UL16_HGMYA	P16757 human cytom	344	17	70.8	302	2	Q8MSM4_ORYSA	Q8MSM4 oryza sativ
272	17	70.8	231	2	Q5TMZ9_ANOGA	Q5TMZ9 anopheles g	345	17	70.8	302	2	Q5FOJ6_GUTOX	Q5FOJ6 glucobact
273	17	70.8	235	2	Q8UTX8_XENLA	Q8UTX8 xenopus lae	346	17	70.8	304	2	Q5TMU6_ANOGA	Q5TMU6 anopheles g
274	17	70.8	236	2	Q5DZAS_VIBFI	Q5DZAS vibrio fusc	347	17	70.8	304	2	Q6M4R9_CORGL	Q6M4R9 corall
275	17	70.8	236	2	Q8UUX7_XENLA	Q8UUX7 xenopus lae	348	17	70.8	305	1	YGBT_ECOLI	Q4XQ66 PLACH
276	17	70.8	237	2	Q4IN40_GIBBE	Q4IN40 gibberella	349	17	70.8	305	2	Q82467_SALTI	Q82467 salmella
277	17	70.8	237	2	Q92XP6_RHIME	Q92XP6 rhizobium m	350	17	70.8	305	2	Q84X66_PLACH	Q84X66 plasmodium
278	17	70.8	238	1	RNRH_RHINI	P08056 rhizopus ni	351	17	70.8	306	2	Q61174_CAEBR	Q61174 caenorhabdi
279	17	70.8	238	2	Q811Z8_CAEBL	Q811Z8 caenorhabdi	352	17	70.8	306	2	Q5TKI6_SALCH	Q5TKI6 salmonella
280	17	70.8	240	2	Q5TSH9_ANOGA	Q5TSH9 anopheles g	353	17	70.8	306	2	Q6LWAS_PROPR	Q6LWAS photobacter
281	17	70.8	240	2	Q4LZX3_9BURK	Q4LZX3 burkholderi	354	17	70.8	306	2	Q8ZMF2_SALTY	Q8ZMF2 geobacter s
282	17	70.8	242	2	Q5TPG2_ANOGA	Q5TPG2 anopheles g	355	17	70.8	307	2	Q74DC4_GEOSL	Q74DC4 geobacter s
283	17	70.8	242	2	Q8EGH3_SHEON	Q8EGH3 shevanelia	356	17	70.8	307	2	Q8X7W3_ECO57	Q8X7W3 escherichia
284	17	70.8	243	1	TRY1_XENLA	P17979 xenopus lae	357	17	70.8	307	2	Q8VGG3_MOUSE	Q8VGG3 mus musculus
285	17	70.8	243	2	Q5M959_XENTR	Q5M959 xenopus tro	358	17	70.8	309	2	Q17234_CAEBL	Q17234 caenorhabdi
286	17	70.8	243	2	Q5M976_XENTR	Q5M976 xenopus tro	359	17	70.8	311	2	Q9V684_DROME	Q9V684 drosophila
287	17	70.8	243	2	Q7S206_XENLA	Q7S206 xenopus lae	360	17	70.8	314	2	Q7RSR0_PLYAO	Q7RSR0 plasmodium
288	17	70.8	243	2	Q8AV83_BRARE	Q8AV83 brachydantio	361	17	70.8	315	2	Q5R3K3_HUMAN	Q5R3K3 homo sapien
289	17	70.8	244	1	KLK6_HUMAN	Q92876 homo sapien	362	17	70.8	315	2	Q97601_RABIT	Q97601 oryctolagus
290	17	70.8	244	2	Q6H301_HUMAN	Q6H301 homo sapien	363	17	70.8	317	2	Q7NVD5_CHRYO	Q7NVD5 chromobacte
291	17	70.8	244	2	Q9XZK6_LETMA	Q9XZK6 leishmania	364	17	70.8	319	2	Q5GGB3_HUMAN	Q5GGB3 homo sapien
292	17	70.8	247	2	Q966B6_CAEBL	Q966B6 caenorhabdi	365	17	70.8	319	2	Q4VSL2_DROME	Q4VSL2 drosophila
293	17	70.8	247	2	Q561Z7_BRARE	Q561Z7 brachydantio	366	17	70.8	319	2	Q4STN3_TETNG	Q4STN3 tetraodon n
294	17	70.8	248	1	CUTC_PORCI	Q7UMB6 porphyromon	367	17	70.8	320	2	Q8NVP3_CORGL	Q8NVP3 corynebacte
295	17	70.8	248	2	Q61HW9_CAEBR	Q61HW9 caenorhabdi	368	17	70.8	326	2	Q4SHX8_TETNG	Q4SHX8 tetraodon n
296	17	70.8	248	2	Q9LZX3_ARATH	Q9LZX3 arabidopsis	369	17	70.8	330	2	Q6MP07_9BURK	Q6MP07 burkholderi
297	17	70.8	250	2	Q6HX41_BACAN	Q6HX41 bacillus an	370	17	70.8	333	2	Q6KX07_XENLA	Q6KX07 xenopus lae
298	17	70.8	251	1	KR10A_HUMAN	P60014 homo sapien	371	17	70.8	334	2	Q4LHW6_9BURK	Q4LHW6 burkholderi
299	17	70.8	252	2	Q8RXT8_ARATH	Q8RXT8 arabidopsis	372	17	70.8	347	2	Q26995_TOXKO	Q26995 toxoplasma
300	17	70.8	253	2	Q5CAQ0_CHICK	Q5CAQ0 gallus gall	373	17	70.8	347	2	Q829C0_STRAW	Q829C0 streptomyce
301	17	70.8	255	1	KR102_HUMAN	P60368 homo sapien	374	17	70.8	348	2	Q9S213_STRTO	Q9S213 streptomyce
302	17	70.8	259	2	Q52T88_VIBCH	Q52T88 vibrio chol	375	17	70.8	350	1	LYM2_ARATH	Q33006 arabidopsis
303	17	70.8	259	2	Q4LY98_9BURK	Q4LY98 burkholderi	376	17	70.8	351	2	Q4S1E9_TETNG	Q4S1E9 tetraodon n
304	17	70.8	260	2	Q7Y150_ORYSA	Q7Y150 oryza sativ	377	17	70.8	352	2	Q5GVU1_XANDR	Q5GVU1 xanthomonas
305	17	70.8	266	2	Q73QF7_TREDE	Q73QF7 treponema d	378	17	70.8	352	2	Q9A6L0_CAUCR	Q9A6L0 caulobacter
306	17	70.8	267	2	Q9A8R0_CAUCR	Q9A8R0 caulobacter	379	17	70.8	358	1	MRAY_NOCPA	Q5YYV2 nocardia fa
307	17	70.8	271	1	KR105_HUMAN	P60370 homo sapien	380	17	70.8	361	2	Q6HH39_BACHK	Q6HH39 bacillus th
308	17	70.8	271	1	MURI_MYCBO	P63635 mycobacteri	381	17	70.8	361	2	Q73SR6_BACCL	Q73SR6 bacillus ce
309	17	70.8	271	1	MURI_MYCTU	P63635 mycobacteri	382	17	70.8	361	2	Q81NW8_BACAN	Q81NW8 bacillus an
310	17	70.8	272	1	MURI_MYCUE	P46705 mycobacteri	383	17	70.8	361	2	Q639G7_BACEZ	Q639G7 bacillus ce
311	17	70.8	275	2	Q73X85_MYCPA	Q73X85 mycobacteri	384	17	70.8	363	2	Q5D8F2_SCHQA	Q5D8F2 schistosoma
312	17	70.8	276	2	Q5MIT5_ABDAL	Q5MIT5 aedes aldob	385	17	70.8	363	2	Q4IWH8_AZOVI	Q4IWH8 azotobacter
313	17	70.8	277	2	Q7MPJ3_VIBAV	Q7MPJ3 vibrio vuln	386	17	70.8	363	2	Q4UXI9_COROK	Q4UXI9 corynebacte
314	17	70.8	278	2	Q7SLH4_ORYSA	Q7SLH4 oryza sativ	387	17	70.8	365	1	KR106_HUMAN	P60371 homo sapien
315	17	70.8	280	2	Q6R8Z3_CTOIN	Q6R8Z3 ctiona intes	388	17	70.8	365	2	Q8FDG2_ECOL6	Q8FDG2 escherichia
316	17	70.8	282	1	KR101_HUMAN	P60331 homo sapien	389	17	70.8	365	2	Q83712_TREPA	Q83712 treponema p
317	17	70.8	283	2	Q4RKP2_TETNG	Q4RKP2 tetraodon n	390	17	70.8	366	2	Q55562_NPVLS	Q55562 leucania se
318	17	70.8	283	2	Q54KZ6_DICDI	Q54KZ6 dictyosteli	391	17	70.8	367	2	Q8NSA6_HUMAN	Q8NSA6 homo sapien
319	17	70.8	283	2	Q4SJTI_TETNG	Q4SJTI tetraodon n	392	17	70.8	367	2	Q6AVJ3_ORYSA	Q6AVJ3 oryza sativ
320	17	70.8	284	2	Q91295_RANCA	Q91295 rana catesb	393	17	70.8	370	1	KR107_HUMAN	P60409 homo sapien
321	17	70.8	286	2	Q4S0J6_TETNG	Q4S0J6 tetraodon n	394	17	70.8	371	2	Q6Z6N5_CAEBR	Q6Z6N5 caenorhabdi
322	17	70.8	290	2	Q7Z729_HUMAN	Q7Z729 homo sapien	395	17	70.8	372	2	Q4RQ43_TETNG	Q4RQ43 tetraodon n
323	17	70.8	290	2	Q6KXZ1_SCYCA	Q6KXZ1 burkholderi	396	17	70.8	372	2	Q63P79_BURPS	Q63P79 burkholderi

397	17	70.8	373	2	Q80YN6_RAT	Q80YN6_rattus norv	470	17	70.8	461	2	Q81ZK8_HUMAN	Q81ZK8_homo sapien
398	17	70.8	375	2	Q7NGJ0_GLOAVI	Q7NGJ0_gloeobacter	471	17	70.8	461	2	Q9B080_HUMAN	Q9B080_homo sapien
399	17	70.8	376	2	Q7NGY1_CLOAB	Q7NGY1_clostridium	472	17	70.8	462	2	Q8MOC6_CAEBL	Q8MOC6_caenorhabdi
400	17	70.8	377	2	Q96V44_TRIRE	Q96V44_trichodema	473	17	70.8	462	2	Q8NOP5_GDELT	Q8NOP5_gnathomys
401	17	70.8	378	2	Q9KJL5_TRIIS1	Q9KJL5_trichodeami	474	17	70.8	467	2	Q584K9_9TRYR	Q584K9_aeareomyscob
402	17	70.8	381	2	Q81G20_CABEL	Q81G20_caenorhabdi	475	17	70.8	467	2	Q17233_BOMMO	Q17233_bombus mori
403	17	70.8	381	2	Q4NZ90_9DELT	Q4NZ90_aeareomyscob	476	17	70.8	467	2	Q4R3M8_MACFA	Q4R3M8_macaca fasc
404	17	70.8	382	2	Q41Q26_9BURK	Q41Q26_burkholderi	477	17	70.8	470	2	Q7QWMS_GIALA	Q7QWMS_giardia lam
405	17	70.8	382	2	Q60C97_METCA	Q60C97_methylococc	478	17	70.8	471	1	VS11_TRYRB	VS11_trypanosoma
406	17	70.8	383	2	Q4SY38_TETNG	Q4SY38_tetradodon n	479	17	70.8	471	2	USMA_TRYRB	USMA_trypanosoma
407	17	70.8	384	2	Q8R7E9_THETN	Q8R7E9_thermoaer	480	17	70.8	471	2	Q5D9E0_SCHJA	Q5D9E0_schistosoma
408	17	70.8	385	2	Q7SEFE_ASHGO	Q7SEFE_aeshbya goss	481	17	70.8	472	2	Q4FYN9_LEIMA	Q4FYN9_leishmania
409	17	70.8	388	2	Q95WA3_ENCIN	Q95WA3_encephalito	482	17	70.8	473	2	Q5XHA0_XENTR	Q5XHA0_xenopus tro
410	17	70.8	388	2	Q5YNC4_NOCFA	Q5YNC4_nocardia s	483	17	70.8	473	2	Q7T0U6_XENLA	Q7T0U6_xenopus lae
411	17	70.8	391	2	Q7ABNO_GEOSL	Q7ABNO_goeobacter f	484	17	70.8	475	2	Q60UR6_CAEBR	Q60UR6_caenorhabdi
412	17	70.8	393	2	Q6AL17_DESPS	Q6AL17_desulfotale	485	17	70.8	475	2	Q61UP8_CAEBR	Q61UP8_caenorhabdi
413	17	70.8	394	2	Q72EB4_DESVH	Q72EB4_desulfovibr	486	17	70.8	476	2	Q9ELS4_HPBVO	Q9ELS4_hepatitis b
414	17	70.8	395	2	Q95V70_TERTH	Q95V70_tetrahymena	487	17	70.8	477	2	Q519Q4_ENTHI	Q519Q4_entamoeba h
415	17	70.8	396	1	IP6K3_MOUSE	Q8BWd2_mus muscucu	488	17	70.8	478	2	Q4W8S1_HUMAN	Q4W8S1_homo sapien
416	17	70.8	396	1	ISCS_RUMFL	Q54055_ruminococcu	489	17	70.8	489	2	Q9W487_DROME	Q9W487_drosophila
417	17	70.8	396	1	Q4LGT2_9BURK	Q4LGT2_burkholderi	490	17	70.8	485	2	Q8S999_PHAAN	Q8S999_phaeolus a
418	17	70.8	397	2	Q95V71_TERTH	Q95V71_tetrahymena	491	17	70.8	486	2	Q58559_9XANT	Q58559_xanthomomas
419	17	70.8	398	1	NIFPS2_ANAVH	Q44482_anabaena va	492	17	70.8	486	2	Q6MRD8_BDEBA	Q6MRD8_bellovibri
420	17	70.8	400	1	NIFS_ANAAZ	Q43884_anabaena az	493	17	70.8	487	2	Q4S723_TETNG	Q4S723_tetradodon n
421	17	70.8	400	1	NIFS_ANASP	P12623_anabaena sp	494	17	70.8	489	1	MDM4_MOUSE	MDM4_mouse
422	17	70.8	400	2	Q09256_SYNP8	Q09256_synechococc	495	17	70.8	489	2	Q9CYG1_MOUSE	Q9CYG1_mouse
423	17	70.8	400	2	Q896F7_CLOTE	Q896F7_clostridium	496	17	70.8	490	1	MDM4_HUMAN	MDM4_human
424	17	70.8	401	1	KR104_HUMAN	P60372_homo sapien	497	17	70.8	490	2	Q6G5T8_HUMAN	Q6G5T8_homo sapien
425	17	70.8	402	2	Q62FW6_BURMA	Q62FW6_burkholderi	498	17	70.8	490	2	Q5XIN1_RAT	Q5XIN1_rattus norv
426	17	70.8	404	2	Q8BLX5_PEEPX	Q8BLX5_pseudomonas	499	17	70.8	490	2	Q99L86_MOUSE	Q99L86_mus muscucu
427	17	70.8	405	2	Q5VKR9_SACER	Q5VKR9_saccharopol	500	17	70.8	492	1	E1BL_ADE07	E1BL_adeno
428	17	70.8	407	2	Q511A1_MAGGR	Q511A1_magnaporthe	501	17	70.8	492	2	Q5EY82_ADE07	Q5EY82_human adeno
429	17	70.8	407	2	Q8B9B2_NPVRO	Q8B9B2_rachiplusia	502	17	70.8	492	2	Q6PBR9_BRARE	Q6PBR9_brycon adeno
430	17	70.8	408	1	VLEN_NPVAC	P24647_autographia	503	17	70.8	494	2	Q571X1_TRYBR	Q571X1_trypanosoma
431	17	70.8	410	1	IP6K3_HUMAN	Q96PC2_homo sapien	504	17	70.8	494	2	Q74BB6_GROSL	Q74BB6_goeobacter s
432	17	70.8	410	2	Q53ER7_HUMAN	Q53ER7_homo sapien	505	17	70.8	498	2	Q4PBP5_USTMA	Q4PBP5_ustilago ma
433	17	70.8	410	2	Q5TAQ4_HUMAN	Q5TAQ4_homo sapien	506	17	70.8	500	2	Q8NA51_HUMAN	Q8NA51_homo sapien
434	17	70.8	410	2	Q5RCO9_PONPY	Q5RCO9_pongo pygma	507	17	70.8	506	2	Q6UD12_9HERP	Q6UD12_giant h
435	17	70.8	410	2	Q7NZY8_CHRVO	Q7NZY8_chromobacte	508	17	70.8	510	2	Q51XM4_MAGGR	Q51XM4_magnaporthe
436	17	70.8	411	2	Q6CGF0_YARLI	Q6CGF0_yarrowia li	509	17	70.8	514	2	Q4IGP6_GIBBE	Q4IGP6_gibberella
437	17	70.8	411	2	Q820TF_ENTPA	Q820TF_enterococcu	510	17	70.8	515	2	Q8H812_TORXA	Q8H812_oryza sativ
438	17	70.8	412	2	Q8YB76_BRUME	Q8YB76_brucella me	511	17	70.8	515	2	Q8UUV9_BRARE	Q8UUV9_brachydantio
439	17	70.8	413	2	Q4QD02_LEIMA	Q4QD02_leishmania	512	17	70.8	517	2	Q5THX6_HUMAN	Q5THX6_homo sapien
440	17	70.8	413	2	Q4RH65_TETNG	Q4RH65_tetradodon n	513	17	70.8	518	2	Q4W8S3_HUMAN	Q4W8S3_homo sapien
441	17	70.8	415	2	Q8GE34_HEIMO	Q8GE34_heliobacilli	514	17	70.8	521	2	Q94718_PARTE	Q94718_parmecium
442	17	70.8	419	2	Q4QDN4_LEIMA	Q4QDN4_leishmania	515	17	70.8	522	2	Q80XU0_MOUSE	Q80XU0_mus muscucu
443	17	70.8	421	2	Q95V69_TERTH	Q95V69_tetrahymena	516	17	70.8	524	1	B028_CAEBL	B028_caenorhabdi
444	17	70.8	423	2	Q8Q0M7_METMA	Q8Q0M7_methanosarc	517	17	70.8	524	2	Q6PFL6_HUMAN	Q6PFL6_homo sapien
445	17	70.8	423	2	Q9U697_TERTH	Q9U697_tetrahymena	518	17	70.8	524	2	Q6OMX6_CAEBR	Q6OMX6_caenorhabdi
446	17	70.8	423	2	Q7XK63_ORYSA	Q7XK63_oryza sativ	519	17	70.8	525	2	Q41507_GIBBE	Q41507_gibberella
447	17	70.8	431	2	Q21650_CAEBL	Q21650_caenorhabdi	520	17	70.8	527	2	Q41815_GIBBE	Q41815_gibberella
448	17	70.8	432	2	Q4Q476_LEIMA	Q4Q476_leishmania	521	17	70.8	530	1	TRAP6_MOUSE	TRAP6_mus muscucu
449	17	70.8	435	2	Q21668_CAEBR	Q21668_caenorhabdi	522	17	70.8	530	2	Q7U6V5_SYNPX	Q7U6V5_synechococ
450	17	70.8	436	2	Q27197_TERTH	Q27197_tetrahymena	523	17	70.8	530	2	Q6P9M0_MOUSE	Q6P9M0_mus muscucu
451	17	70.8	437	2	Q5JLKS_ORYSA	Q5JLKS_oryza sativ	524	17	70.8	532	2	Q9SKS0_ARATH	Q9SKS0_arabidopsis
452	17	70.8	438	2	Q8AXV2_FUGRU	Q8AXV2_fugu rubrip	525	17	70.8	542	2	Q7R0Z0_GIALA	Q7R0Z0_giardia lam
453	17	70.8	439	2	Q94827_TERTH	Q94827_tetrahymena	526	17	70.8	543	2	Q9UVL7_PNECA	Q9UVL7_pneumocysti
454	17	70.8	439	2	Q8JID1_BRARE	Q8JID1_brachydantio	527	17	70.8	543	2	Q4P677_USTMA	Q4P677_ustilago ma
455	17	70.8	440	2	Q8N1H1_HUMAN	Q8N1H1_homo sapien	528	17	70.8	546	2	Q68EW3_XENLA	Q68EW3_xenopus lae
456	17	70.8	444	2	Q4QBT5_LEIMA	Q4QBT5_leishmania	529	17	70.8	547	1	C78A1_MAI2E	C78A1_maize
457	17	70.8	445	2	Q9NC96_9TRYR	Q9NC96_typanosoma	530	17	70.8	547	2	Q41295_GIBBE	Q41295_gibberella
458	17	70.8	445	2	Q57X19_9TRYR	Q57X19_typanosoma	531	17	70.8	547	2	Q4T338_TETNG	Q4T338_tetradodon n
459	17	70.8	447	2	Q96FD7_HUMAN	Q96FD7_homo sapien	532	17	70.8	549	2	Q96SV9_HUMAN	Q96SV9_homo sapien
460	17	70.8	447	2	Q8PLV4_XANAC	Q8PLV4_xanthomomas	533	17	70.8	553	2	Q6BP88_PARTE	Q6BP88_parmecium
461	17	70.8	449	2	Q53RR6_HUMAN	Q53RR6_homo sapien	534	17	70.8	554	2	Q5QJH1_TIG	Q5QJH1_tus scrofa
462	17	70.8	453	2	Q51P64_MAGGR	Q51P64_magnaporthe	535	17	70.8	555	1	SYT14_HUMAN	SYT14_homo sapien
463	17	70.8	454	2	Q8C1S8_MOUSE	Q8C1S8_mus muscucu	536	17	70.8	555	2	YL133_CAEBL	YL133_caenorhabdi
464	17	70.8	456	1	RINI_RAT	P29315_rattus norv	537	17	70.8	555	2	Q5THX7_HUMAN	Q5THX7_homo sapien
465	17	70.8	456	2	Q924P4_MOUSE	Q924P4_mus muscucu	538	17	70.8	556	2	Q6DUJ2_XENLA	Q6DUJ2_xenopus lae
466	17	70.8	456	2	Q91V17_MOUSE	Q91V17_mus muscucu	539	17	70.8	563	2	Q5MRV3_LEGFL	Q5MRV3_legionella
467	17	70.8	456	2	Q6IRS9_RAT	Q6IRS9_rattus norv	540	17	70.8	567	2	Q9LXC1_ARATH	Q9LXC1_arabidopsis
468	17	70.8	460	1	RINI_HUMAN	P13489_homo sapien	541	17	70.8	574	2	Q5THX8_HUMAN	Q5THX8_homo sapien
469	17	70.8	460	2	Q871X0_NEUCR	Q871X0_neurospora	542	17	70.8	577	1	ITB6_CAVPO	ITB6_cavia porce

543	17	70.8	582	2	062U00_BURMA	06cju0 burkholderi	616	17	70.8	796	2	05T9V1_HUMAN	05t9v1 homo sapien
544	17	70.8	585	2	09IF49_HPBVO	09if49 hepatitis b	617	17	70.8	807	2	05T9V6_9TRYP	05t9v6 trypanosoma
545	17	70.8	596	2	05VT17_HUMAN	05vt17 homo sapien	618	17	70.8	812	2	06T6B3_GALLI	06t6b3 gallus galli
546	17	70.8	596	2	08T6A8_CAEEL	08t6a8 caenorhabdi	619	17	70.8	813	2	05DVW0_AERBO	05dvw0 aeromonas s
547	17	70.8	596	2	06T1R2_RAT	06t1r2 rattus norv	620	17	70.8	814	2	080VY4_MOUSE	080vy4 mus musculu
548	17	70.8	601	2	08LM37_ORYSA	08lm37 oryza sativ	621	17	70.8	820	2	04SUN7_TETNG	04sun7 tetradon n
549	17	70.8	602	2	04P939_USTMA	04p939 uetillago ma	622	17	70.8	823	2	04SUN7_TETNG	04sun7 tetradon n
550	17	70.8	603	2	06ZJR2_ORYSA	06zjr2 oryza sativ	623	17	70.8	823	2	04SUN7_TETNG	04sun7 tetradon n
551	17	70.8	603	2	04NNZ1_9DELT	04nnz1 anaeromyxob	624	17	70.8	833	2	06J288_ACACA	06j288 acanthamoeb
552	17	70.8	604	2	09NXGS_HUMAN	09nxgs homo sapien	625	17	70.8	836	2	08UJ66_HPBVO	08uj66 hepatitis b
553	17	70.8	608	2	07YF66_CRYPV	07yf66 cryptospori	626	17	70.8	843	1	DBOL_HPBVI	dbol hepatitis b
554	17	70.8	609	2	05ZGAI_MAGR	05zgai magnaporthe	627	17	70.8	843	2	068R57_HPBVO	068r57 hepatitis b
555	17	70.8	609	2	05CPJ1_CRYPV	05cpj1 cryptospori	628	17	70.8	843	2	080H36_HPBVO	080h36 hepatitis b
556	17	70.8	612	2	06JUV6_BURPS	06juv6 burkholderi	629	17	70.8	843	2	09QMJ3_HPBVO	09qmj3 hepatitis b
557	17	70.8	612	2	04SBD5_TETNG	04sbd5 tetradon n	630	17	70.8	843	2	09QMJ3_HPBVO	09qmj3 hepatitis b
558	17	70.8	613	2	04N940_THEPA	04n940 theileria p	631	17	70.8	843	2	09YL91_HPBVO	09yl91 hepatitis b
559	17	70.8	615	2	09LOP9_ARATH	09lop9 arabidopsis	632	17	70.8	843	2	003766_HPBVO	003766 hepatitis b
560	17	70.8	618	2	07S5R7_NEUCR	07s5r7 neurospora	633	17	70.8	844	2	05H116_XANOR	05h116 xanthomonas
561	17	70.8	629	2	07ZTW7_BRARE	07ztw7 brachydantio	634	17	70.8	849	2	04PEQ2_USTMA	04peq2 uetillago ma
562	17	70.8	631	2	048523_ARATH	048523 arabidopsis	635	17	70.8	849	2	05YQX9_NOCFA	05yqx9 nocardia fa
563	17	70.8	638	2	09LOP8_ARATH	09lop8 arabidopsis	636	17	70.8	853	2	08VKP5_MYCTU	08vkp5 mycobacteri
564	17	70.8	640	2	09N520_CAEEL	09n520 caenorhabdi	637	17	70.8	853	2	07U2G2_MYCBO	07u2g2 mycobacteri
565	17	70.8	650	2	058B38_BRARE	058b38 brachydantio	638	17	70.8	853	2	053674_MYCTU	053674 mycobacteri
566	17	70.8	653	2	04KMG1_HUMAN	04kmg1 homo sapien	639	17	70.8	854	2	05ZBC9_TORYSA	05zbc9 oryza sativ
567	17	70.8	654	2	08QZZ0_MOUSE	08qzz0 mus musculu	640	17	70.8	855	2	06GVN7_MYCAV	06gvn7 mycobacteri
568	17	70.8	655	2	09H0T6_HUMAN	09h0t6 homo sapien	641	17	70.8	855	2	07JTL6_MYCPA	07jtl6 mycobacteri
569	17	70.8	656	2	054RH5_DICDI	054rh5 dictyosteli	642	17	70.8	855	2	06FB25_ACIAAD	06fb25 actinobact
570	17	70.8	656	2	05U227_RAT	05u227 rattus norv	643	17	70.8	864	2	04ST29_TETNG	04st29 tetradon n
571	17	70.8	658	2	018215_CAEEL	018215 caenorhabdi	644	17	70.8	870	2	06PCD4_HUMAN	06pcd4 homo sapien
572	17	70.8	659	2	05ZCQ0_ORYSA	05zcg0 oryza sativ	645	17	70.8	873	2	04S9L8_TETNG	04s9l8 tetradon n
573	17	70.8	663	2	052D83_MAGGR	052d83 magnaporthe	646	17	70.8	879	2	09XZ14_DROME	09xz14 mus musculu
574	17	70.8	667	1	TS11_GIALA	003185 giardia lam	647	17	70.8	893	1	RUSC1_MOUSE	08NC14_HUMAN
575	17	70.8	667	2	0627F3_CABRR	0627f3 caenorhabdi	648	17	70.8	894	2	05RER7_PONPY	05rer7 ponpy
576	17	70.8	670	2	077423_BACRY	077423 brachydantio	649	17	70.8	894	2	05XFR8_CRYNE	05xfr8 cryptococcu
577	17	70.8	674	2	04P1N4_USTMA	04p1n4 uetillago ma	650	17	70.8	900	2	05XFR8_CRYNE	05xfr8 cryptococcu
578	17	70.8	676	2	052K08_XENLA	052k08 xenopus lae	651	17	70.8	900	2	05XFR8_CRYNE	05xfr8 cryptococcu
579	17	70.8	677	1	051W16_9REOV	051w16 avian ortho	652	17	70.8	902	1	RUSC1_HUMAN	05t9v2_HUMAN
580	17	70.8	677	1	WHITE_LUCCU	005360 lucilia cup	653	17	70.8	902	2	05T9V2_HUMAN	05t9v2_HUMAN
581	17	70.8	679	2	081S30_BACCU	081s30 bactrocera	654	17	70.8	903	2	055XB7_CRYNE	055xb7 cryptococcu
582	17	70.8	681	2	07ZZ29_BRARE	07zz29 brachydantio	655	17	70.8	904	2	05KML6_CRYNE	05kml6 cryptococcu
583	17	70.8	682	2	07YQK2_BOVIN	07yqk2 bos taurus	656	17	70.8	904	2	TLR3_HUMAN	0504W0_HUMAN
584	17	70.8	684	2	081498_CUPSA	081498 cuplemnius	657	17	70.8	904	2	0504W0_HUMAN	0504w0_HUMAN
585	17	70.8	684	2	07ZVU9_BRARE	07zvu9 brachydantio	658	17	70.8	904	2	04VAL2_HUMAN	04val2_HUMAN
586	17	70.8	687	1	WHITE_DROME	PI0090 drosophila	659	17	70.8	909	2	06MP91_BDEBA	06mp91 belliovlbri
587	17	70.8	687	1	094960_DROSI	094960 drosophila	660	17	70.8	912	2	055N08_CRYNE	055n08 cryptococcu
588	17	70.8	691	2	04Q0B7_LEIMA	04q0b7 leishmania	661	17	70.8	912	2	05KCA3_CRYNE	05kca3 cryptococcu
589	17	70.8	695	2	07LSY4_HUMAN	07lsy4 homo sapien	662	17	70.8	921	2	04UGJ3_THEAN	04ugj3 theileria a
590	17	70.8	695	2	058D07_BOVIN	058d07 bos taurus	663	17	70.8	929	2	05TVH8_9TRYP	05tvh8 trypanosoma
591	17	70.8	705	2	04T7V1_TETNG	04t7v1 tetradon n	664	17	70.8	934	1	CO6_HUMAN	05TVH8_9TRYP
592	17	70.8	707	2	06C4V7_VARLI	06c4v7 yarrowia li	665	17	70.8	934	1	CO6_PANTR	05TVH8_9TRYP
593	17	70.8	707	2	08BM61_MOUSE	08bm61 mus muscu	666	17	70.8	934	1	CO6_PANTR	05TVH8_9TRYP
594	17	70.8	712	2	04FX57_LEIMA	04fx57 leishmania	667	17	70.8	938	2	07YTA2_CRYPV	07yta2 cryptospori
595	17	70.8	715	2	05RD19_PONPY	05rd19 pongo pygma	668	17	70.8	938	2	05CLY0_CRIHO	05cly0 cryptospori
596	17	70.8	719	2	09U019_GIALA	09u019 giardia lam	669	17	70.8	938	2	04K9C6_PSEB5	04k9c6 pseudomonas
597	17	70.8	725	2	09U021_GIALA	09u021 giardia lam	670	17	70.8	950	2	08RZK0_ORYSA	08rzko oryza sativ
598	17	70.8	719	2	07YOK1_BOVIN	07yok1 bos taurus	671	17	70.8	969	2	07R604_GIALA	07r604 giardia lam
599	17	70.8	725	2	08BZW8_MOUSE	08bzw8 mus musculu	672	17	70.8	975	2	08NCC9_HUMAN	08ncc9 homo sapien
600	17	70.8	726	2	08NBZ2_HUMAN	08nbz2 homo sapien	673	17	70.8	983	2	05HBA4_HUMAN	05hba4 homo sapien
601	17	70.8	737	2	08BW83_MOUSE	08bw83 m mus muscu	674	17	70.8	993	2	04QUG9_LEIMA	04qug9 leishmania
602	17	70.8	746	2	08S6L6_ORYSA	08s6l6 oryza sativ	675	17	70.8	999	2	04SQC0_TETNG	04sqc0 tetradon n
603	17	70.8	747	2	08C4V1_MOUSE	08c4v1 mus musculu	676	17	70.8	1002	2	09SMA4_ENCIN	09sma4 encornallito
604	17	70.8	747	2	06NM57_BRARE	06nm57 brachydantio	677	17	70.8	1003	2	08IMQ3_DROME	08imq3 drosophila
605	17	70.8	748	2	08N264_BURAR	08n264 burkholderi	678	17	70.8	1011	2	04SOD2_TETNG	04sod2 tetradon n
606	17	70.8	748	2	09CWM6_MOUSE	09cwm6 mus musculu	679	17	70.8	1021	2	04TB22_TETNG	04tb22 tetradon n
607	17	70.8	749	2	04LT50_9BURK	04lt50 burkholderi	680	17	70.8	1023	2	08SB11_ORYSA	08sb11 oryza sativ
608	17	70.8	751	2	08LRK4_ARATH	08lrk4 arabidopsis	681	17	70.8	1023	2	07XGB0_ORYSA	07xgb0 oryza sativ
609	17	70.8	751	2	04M064_9BURK	04m064 burkholderi	682	17	70.8	1031	2	076836_CAEEL	076836 caenorhabdi
610	17	70.8	769	2	04P3X8_USTMA	04p3x8 uetillago ma	683	17	70.8	1049	2	06JUN47_LYCES	06jun47 lycopersico
611	17	70.8	783	2	05MT26_LEGPI	05mt26 legionella	684	17	70.8	1059	2	04Y187_PLACH	04y187 plasmidium
612	17	70.8	783	2	05X1B3_LEGPA	05x1b3 legionella	685	17	70.8	1065	1	LRIG2_HUMAN	08chf0 mus musculu
613	17	70.8	788	1	ITB6_HUMAN	PI5564 homo sapien	686	17	70.8	1080	1	HDC_DROME	094898 homo sapien
614	17	70.8	794	2	09U110_LEIMA	09u110 leishmania	687	17	70.8	1089	2	08T3A0_CIOIN	08t3a0 drosophila
615	17	70.8					688	17	70.8				08t3a0 ciona intes



689	17	70.8	1091	2	Q4T4A4_TETNG	Q4c4a4 tetradon n	762	17	70.8	3063	1	COCA1_HUMAN	Q99715 homo sapien
690	17	70.8	1107	2	Q6MKF5_BDEBA	Q6mkf5 bdellovibri	763	17	70.8	3063	1	COCA1_HUMAN	Q99715 homo sapien
691	17	70.8	1109	2	Q61YU2_CAEBR	Q61yu2 caenorhabdi	764	17	70.8	3119	2	COCA1_MOUSE	Q60847 mus musculu
692	17	70.8	1156	1	PMP2_CAEEL	P2p192 caenorhabdi	765	17	70.8	3124	1	COCA1_CHICK	P139944 gallus gall
693	17	70.8	1183	2	Q5DU14_MOUSE	Q5du14 mus musculu	766	17	70.8	3229	2	Q26912_TRYCR	Q26912 trypanosoma
694	17	70.8	1207	2	Q8QNC2_9PHYC	Q8qnc2 eucoccarpus	767	17	70.8	3457	2	Q6EEU1_VIVIRU	Q6eeu1 maize chlor
695	17	70.8	1209	2	Q9SGS6_ARATH	Q9sgs6 arabidopsis	768	17	70.8	3616	1	RFOA_LIDVP	Q83017 l replicase
696	17	70.8	1232	2	Q4T1X3_TETNG	Q4t1x3 tetradon n	769	17	70.8	3881	2	Q4ZG84_HUMAN	Q4zg84 homo sapien
697	17	70.8	1236	2	Q3NED7_LEIMA	Q3ned7 leishmania	770	17	70.8	3982	2	Q4RV74_TETNG	Q4rv74 tetradon n
698	17	70.8	1262	1	DXH9_XENLA	Q66fk8 xenopus lae	771	17	70.8	4655	1	LRP2_HUMAN	P81664 homo sapien
699	17	70.8	1269	1	DXH9_PONPY	Q58t74 pongo pygma	772	17	70.8	4655	2	Q7Z5C0_HUMAN	Q7z5c0 homo sapien
700	17	70.8	1270	1	DXH9_HUMAN	Q08211 homo sapien	773	17	70.8	4655	2	Q7Z5C1_HUMAN	Q7z5c1 homo sapien
701	17	70.8	1270	2	Q4PGU2_USTMA	Q4pgu2 ustilago ma	774	17	70.8	4856	2	Q61EJ2_CAEBR	Q61ej2 caenorhabdi
702	17	70.8	1287	1	DXH9_BOVIN	Q28141 bos taurus	775	17	70.8	4936	2	Q6D5C2_ERWCT	Q6d5c2 erwilia car
703	17	70.8	1323	1	PASK_HUMAN	Q6gsj5 homo sapien	776	17	66.7	24	2	Q71UR1_CHICK	Q71ur1 gallus gall
704	17	70.8	1323	2	Q6GSJ5_HUMAN	Q6gsj5 homo sapien	777	17	66.7	31	2	Q5C4I4_SCHJA	Q5c4i4 schistosoma
705	17	70.8	1323	2	Q6BDY3_HUMAN	Q6bdy3 homo sapien	778	17	66.7	37	2	Q6EAS9_SENTR	Q6eas9 salmonella
706	17	70.8	1327	1	TKNS1_HUMAN	Q95271 homo sapien	779	17	66.7	38	2	Q06516_ECOLI	Q06516 escherichia
707	17	70.8	1327	2	Q4G0F2_HUMAN	Q4g0f2 homo sapien	780	17	66.7	38	2	Q6EAT4_SENTR	Q6eat4 salmonella
708	17	70.8	1345	2	Q66675_9GAMA	Q66675 equid herpe	781	17	66.7	38	2	Q9XBP2_ECOLI	Q9xbp2 escherichia
709	17	70.8	1350	2	Q9VSJ7_DROME	Q9vsj7 drosophila	782	17	66.7	38	2	Q53559_ECOLI	Q53559 escherichia
710	17	70.8	1368	2	Q52E29_MAGGR	Q52e29 magnaporthe	783	17	66.7	39	1	MER11_EUPRA	P26887 euploetes ra
711	17	70.8	1370	2	Q7KXK6_DROME	Q7kqx6 drosophila	784	17	66.7	41	1	AMP_IFONI	P81591 ipomeea nil
712	17	70.8	1377	2	Q4SVT4_TETNG	Q4svt4 tetradon n	785	17	66.7	43	2	Q62C07_BURMA	Q62c07 burkholderi
713	17	70.8	1380	1	DXH9_MOUSE	Q70133 mus musculu	786	17	66.7	44	1	TXPL2_PLETR	P34079 plectreureys
714	17	70.8	1408	2	Q4H229_GIBZE	Q4h229 gibberella	787	17	66.7	44	2	Q9UP90_HUMAN	Q9up90 homo sapien
715	17	70.8	1427	2	Q8VIB7_MESAU	Q8vib7 mesocricetu	788	17	66.7	44	2	Q9TXE7_CAEEL	Q9txe7 caenorhabdi
716	17	70.8	1428	1	ATRN_MOUSE	Q9u460 mus musculu	789	17	66.7	44	2	Q4V648_DROME	Q4v648 drosophila
717	17	70.8	1432	1	ATRN_RAT	Q99j86 rattus norv	790	17	66.7	44	2	Q8G001_BRUSU	Q8g001 brucella bu
718	17	70.8	1442	2	Q5CMW4_CRYHO	Q5cmw4 cryptospori	791	17	66.7	45	1	FULC_MYXFU	P01547 myxococcus
719	17	70.8	1443	2	Q4Q3F5_LEIMA	Q4q3f5 leishmania	792	17	66.7	45	2	Q9N1T1_ORNAN	Q9n1t1 ornithiohyu
720	17	70.8	1445	2	Q5CUC9_CRYPV	Q5cuc9 cryptospori	793	17	66.7	46	2	Q8NH55_HUMAN	Q8nh55 homo sapien
721	17	70.8	1446	2	Q4SDH3_TETNG	Q4sdh3 tetradon n	794	17	66.7	46	2	Q96RW9_HUMAN	Q96rw9 homo sapien
722	17	70.8	1450	2	Q4RP14_TETNG	Q4rp14 tetradon n	795	17	66.7	47	2	Q7OZT7_GIALA	Q7ozt7 giardia lam
723	17	70.8	1457	2	Q8IS20_DICDI	Q8is20 dictyosteli	796	17	66.7	48	2	Q7S4T0_NEUCR	Q7s4t0 neurospora
724	17	70.8	1467	2	Q6IQN6_CAEBR	Q6iqn6 caenorhabdi	797	17	66.7	50	2	Q5DQ60_ECOLI	Q5dq60 escherichia
725	17	70.8	1474	2	Q62504_CAEEL	Q62504 caenorhabdi	798	17	66.7	50	2	Q9DWN9_9PICO	Q9dwn9 hepatitis a
726	17	70.8	1486	2	Q4SDM0_TETNG	Q4sdm0 tetradon n	799	17	66.7	51	2	Q57JF2_SALCH	Q57jf2 salmonella
727	17	70.8	1506	2	Q5KTC0_MOUSE	Q5ktc0 mus musculu	800	17	66.7	52	2	Q7QSV6_GIALA	Q7qsv6 giardia lam
728	17	70.8	1507	1	TRPM2_MOUSE	Q91y44 mus musculu	801	17	66.7	55	1	NXB2_CERLA	P01556 cerebratulu
729	17	70.8	1508	2	Q5G856_RAT	Q5g856 rattus norv	802	17	66.7	55	2	Q9P9Y7_XYLPA	Q9p9y7 xyloella fas
730	17	70.8	1514	2	Q4UAD0_THEAN	Q4uad0 theileria a	803	17	66.7	57	1	GRN1_CYPCA	P81013 cyprinus ca
731	17	70.8	1550	2	Q9LGI1_ARATH	Q9lgi1 arabidopsis	804	17	66.7	57	1	GRN2_CYPCA	P81014 cyprinus ca
732	17	70.8	1557	2	Q4SH63_TETNG	Q4sh63 tetradon n	805	17	66.7	57	1	GRN3_CYPCA	P81015 cyprinus ca
733	17	70.8	1644	2	Q4FX35_LEIMA	Q4fx35 leishmania	806	17	66.7	57	2	Q4LBG5_HUMAN	Q4lbg5 homo sapien
734	17	70.8	1643	2	Q25312_LEIMA	Q25312 leishmania	807	17	66.7	57	2	Q9PRN7_CARAU	Q9prn7 caesarius a
735	17	70.8	1687	1	PHLP_MOUSE	Q8eh44 mus musculu	808	17	66.7	58	2	Q4LBG4_HUMAN	Q4lbg4 gluconobac
736	17	70.8	1696	1	PHLP_RAT	Q9wct8 rattus norv	809	17	66.7	60	2	Q5FSI4_GLUOX	Q5fsi4 gluconobac
737	17	70.8	1717	1	PHLP_HUMAN	Q60346 homo sapien	810	17	66.7	60	2	Q01304_TREPA	Q01304 treponema p
738	17	70.8	1748	1	Q6ITP1_CAEBR	Q6itp1 caenorhabdi	811	17	66.7	61	2	Q7UMI1_RHOBA	Q7umi1 rhodopirell
739	17	70.8	1796	1	COEAL_HUMAN	Q05707 homo sapien	812	17	66.7	62	2	Q4WN22_ASPTU	Q4wn22 aspergillus
740	17	70.8	1797	1	COEAL_MOUSE	Q80x19 mus musculu	813	17	66.7	62	1	MT1K_HUMAN	P80236 homo sapien
741	17	70.8	1866	2	Q4GYP7_9TRYP	Q4gyp7 trypanosoma	814	17	66.7	62	1	MT2A_RABIT	P81805 oryctolagus
742	17	70.8	1888	1	COEAL_CHICK	P32018 gallus gall	815	17	66.7	62	1	MT2C_RABIT	P80230 oryctolagus
743	17	70.8	1917	2	Q515W3_ENTHI	Q515w3 entamoeba h	816	17	66.7	63	2	Q7UND1_RHOBA	Q7und1 rhodopirell
744	17	70.8	1972	2	Q4QW8_LEIMA	Q4qw8 leishmania	817	17	66.7	63	2	Q23803_CHITE	Q23803 chironomus
745	17	70.8	1984	2	Q7RBL5_PLAYO	Q7rbl5 plasmodium	818	17	66.7	65	2	Q8MK64_SHEEP	Q8mk64 ovis aries
746	17	70.8	2083	2	Q8MK9_LEIMA	Q8mk9 leishmania	819	17	66.7	65	2	Q5GM87_XANOR	Q5gm87 xanthomonas
747	17	70.8	2112	2	Q7PD7_PLAYO	Q7pd7 plasmodium	820	17	66.7	66	2	Q9N625_CONCT	Q9n625 conus catu
748	17	70.8	2128	2	Q5ONG8_ENTHI	Q5ong8 entamoeba h	821	17	66.7	66	2	Q9N628_CONCT	Q9n628 conus catu
749	17	70.8	2189	2	Q9B105_EIMTE	Q9b105 elmetria ten	822	17	66.7	66	2	Q9NCV6_CONCT	Q9ncv6 conus catu
750	17	70.8	2204	2	Q8IL70_PLAF7	Q8il70 plasmodium	823	17	66.7	66	2	Q9NCM1_CONCT	Q9ncm1 conus catu
751	17	70.8	2225	2	Q4SXE3_TETNG	Q4sxe3 tetradon n	824	17	66.7	66	2	Q9NCW2_CONCT	Q9ncw2 conus catu
752	17	70.8	2378	2	Q8I3U0_PLAF7	Q8i3u0 plasmodium	825	17	66.7	68	2	Q9U157_HUMAN	Q9u157 homo sapien
753	17	70.8	2395	2	Q27167_PARTE	Q27167 paramecium	826	17	66.7	68	2	Q7YVQ0_9TRYP	Q7yvq0 trypanosoma
754	17	70.8	2408	2	Q57UP3_9TRYP	Q57up3 trypanosoma	827	17	66.7	68	2	Q5WM83_AEDAE	Q5wm83 aedes aegyp
755	17	70.8	2519	2	Q51WN3_MAGGR	Q51wn3 magnaporthe	828	17	66.7	69	2	Q9N6A4_9CAEN	Q9n6a4 conus ebrae
756	17	70.8	2511	2	Q4NKK5_9DELT	Q4nkk5 anaeromyxob	829	17	66.7	69	2	Q9NCT0_9CAEN	Q9nc0 conus ebrae
757	17	70.8	2533	2	P90589_PARTE	P90589 paramecium	830	17	66.7	69	2	Q9NCT1_9CAEN	Q9nc1 conus ebrae
758	17	70.8	2533	2	Q27183_PARTE	Q27183 paramecium	831	17	66.7	69	2	Q9NCT4_9CAEN	Q9nc4 conus ebrae
759	17	70.8	2828	2	Q4STC7_TETNG	Q4stc7 tetradon n	832	17	66.7	69	2	Q9NCT5_9CAEN	Q9nc5 conus ebrae
760	17	70.8	2878	2	Q5CUB1_CRYPV	Q5cub1 cryptospori	833	17	66.7	69	2	Q9NCT6_9CAEN	Q9nc6 conus ebrae
761	17	70.8	2884	2	Q5VYK2_HUMAN	Q5vyk2 homo sapien	834	17	66.7	69	2	Q9NCT7_9CAEN	Q9nc7 conus ebrae





```
981 16 66.7 122 2 063FJ9_BAC2Z 063fj9 bacillus ce
982 16 66.7 122 2 05UVZ1_9ADEN 05uvz1 human adeno
983 16 66.7 122 2 077923_ADEB35 077923 human adeno
984 16 66.7 122 2 08B850_9ADEN 08b850 human adeno
985 16 66.7 123 1 11Z1_PBRMA 080xg2 petromyscus
986 16 66.7 123 2 07P33_ANOGA 07p33 anopheles g
987 16 66.7 123 2 08JG94_GINCI 08jg94 ginglymosto
988 16 66.7 124 2 08PV06_METMA 08pv06 methamosarc
989 16 66.7 124 2 07R266_GIALA 07r266 giardia lam
990 16 66.7 124 2 016121_TENMO 016121 tenebrio mo
991 16 66.7 124 2 062710_CAEBR 062710 caenorhabdi
992 16 66.7 124 2 09U744_TENMO 09u744 tenebrio mo
993 16 66.7 125 2 06DLX5_TENMO 06dlx5 tenebrio mo
994 16 66.7 125 2 0814J7_CABEL 0814j7 caenorhabdi
995 16 66.7 125 2 038346_BPLH 038346 lactococcus
996 16 66.7 125 2 04NR72_9DELT 04nr72 anaeromyxob
997 16 66.7 125 2 08JF04_GINCI 08jf04 ginglymosto
998 16 66.7 125 2 08JF07_GINCI 08jf07 ginglymosto
999 16 66.7 125 2 08JF08_GINCI 08jf08 ginglymosto
1000 16 66.7 125 2 08JFK0_GINCI 08jfk0 ginglymosto
```

## ALIGNMENTS

```
RESULT 1
Q9N669_9CAEN PRELIMINARY; PRT; 69 AA.
AC Q9N669_9CAEN PRELIMINARY; PRT; 69 AA.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Four-loop conotoxin (Fragment).
OS Conus ebraeus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeodonta; Hypogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=89425;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20416305; PubMed=10958845;
RA Duda T.F., Palumbi S.R.;
RT "Evolutionary diversification of multigene families: allelic selection
of toxins in predatory cone snails."
RL Mol. Biol. Evol. 17:1286-1293(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Duda T.F., Palumbi S.R.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF174281; AAF89945.1; -; mRNA.
DR EMBL; AF174280; AAF89944.1; -; mRNA.
DR GO: GO:0005576; C:extracellular region; IEA.
DR GO: GO:0008200; F:ion channel inhibitor activity; IEA.
DR InterPro; IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
FT NON TER 1
SQ SEQUENCE 69 AA; 7662 MW; C8C826FDC920C4F1 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 69;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Four-loop conotoxin (Fragment).
OS Conus ebraeus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeodonta; Hypogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=89425;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20416305; PubMed=10958845;
RA Duda T.F., Palumbi S.R.;
RT "Evolutionary diversification of multigene families: allelic selection
of toxins in predatory cone snails."
RL Mol. Biol. Evol. 17:1286-1293(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Duda T.F., Palumbi S.R.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF174287; AAF89951.1; -; mRNA.
DR GO: GO:0005576; C:extracellular region; IEA.
DR GO: GO:0008200; F:ion channel inhibitor activity; IEA.
DR GO: GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
FT NON TER 1
SQ SEQUENCE 69 AA; 7674 MW; 83D526FDC934422A CRC64;
```

```
Query Match 75.0%; Score 18; DB 2; Length 69;
Best Local Similarity 25.0%; Pred. No. 2e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXX 8
Db 61 CSTATSTC 68

RESULT 3
Q9NCS9_9CAEN PRELIMINARY; PRT; 69 AA.
AC Q9NCS9_9CAEN PRELIMINARY; PRT; 69 AA.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Four-loop conotoxin (Fragment).
OS Conus ebraeus.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeodonta; Hypogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=89425;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20416305; PubMed=10958845;
RA Duda T.F., Palumbi S.R.;
RT "Evolutionary diversification of multigene families: allelic selection
of toxins in predatory cone snails."
RL Mol. Biol. Evol. 17:1286-1293(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Duda T.F., Palumbi S.R.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF174286; AAF89950.1; -; mRNA.
DR GO: GO:0005576; C:extracellular region; IEA.
DR GO: GO:0008200; F:ion channel inhibitor activity; IEA.
DR GO: GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR004214; Conotoxin.
DR Pfam; PF02950; Conotoxin; 1.
FT NON TER 1
SQ SEQUENCE 69 AA; 7650 MW; DF3826FDC920C4F1 CRC64;
```





AC Q82854;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE TAT protein (Transactivating regulatory protein).  
 GN Name=tat;  
 OS Jembrana disease virus.  
 OS Viruses; Retro-transcribing viruses; Retroviridae; Orthoretrovirinae;  
 CC Lentivirus; Bovine lentiviruses.  
 CC NCB1\_TaxID=36370;  
 OK NCB1\_TaxID=36370;  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=Tabanan/87;  
 RX MEDLINE=97201584; PubMed=9049370;  
 RA Chadwick B.J., Coelen R.J., Wilcox G.E., Samuels L.M., Kertyadnya G.,  
 RT "Nucleotide sequence analysis of Jembrana disease virus: a bovine  
 RT lentivirus associated with an acute disease syndrome.";  
 RL J. Gen. Virol. 76:1637-1650(1995).  
 DR EMBL; U21603; AA64395.1; -; Genomic RNA.  
 DR GO; GO:0042025; C:host cell nucleus; IEA.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR GO; GO:0006350; P:transcription; IEA.  
 DR InterPro; IPR001831; IV\_Tat.  
 DR Pfam; PF00539; Tat, 1.  
 KM Activator; Nuclear protein; RNA-binding; Transcription;  
 KM Transcription regulation.  
 SQ SEQUENCE 114 AA; 12457 MW; 77A7CB6FAF128D5A CRC64;

Query Match 75.0%; Score 18; DB 2; Length 114;  
 Best Local Similarity 25.0%; Pred. No. 2.5e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
 Db 96 CASSASC 103

RESULT 10  
 Q9EOL8 MOUSE PRELIMINARY; PRT; 161 AA.  
 AC Q9EOL8;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE T1S1LD deletion variant (Fragment).  
 GN Name=zfp3612; Synonyms=Brlf2;  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 CC Muridae; Murinae; Mus.  
 CC NCB1\_TaxID=10090;  
 OK NCB1\_TaxID=10090;  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=C57BLKS/J;  
 RA Cho K., Hobson K., Greenhalgh D.G.;  
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF205222; AAG45250.1; -; mRNA.  
 DR HSSP; P22893; IMGO.  
 DR SMR; Q9EOL8; 1-49.  
 DR MGI; MGI:107945; Zfp3612.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR000571; Znf\_CCH.  
 DR Pfam; PF00642; zf-CCH; 1.  
 DR SMART; SMO0356; Znf\_C3H1; 1.  
 FT NON TER 1 1  
 KM NON TER 161 161  
 SQ SEQUENCE 161 AA; 17101 MW; BF09DF73709215EE CRC64;

Query Match 75.0%; Score 18; DB 2; Length 161;  
 Best Local Similarity 25.0%; Pred. No. 2.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 1 CXXXXXXC 8  
 Db 129 CASSASC 136

RESULT 11  
 Q7OPW4 GIALA PRELIMINARY; PRT; 163 AA.  
 ID Q7OPW4;  
 AC Q7OPW4;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE GLP 54.16521.16030.  
 OS Giardia lamblia ATCC 50803.  
 CC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
 CC NCB1\_TaxID=184922;  
 OK NCB1\_TaxID=184922;  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=WB C6;  
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,  
 RA Olsen G.J., Sogin M.L.;  
 RT "Draft sequence of the Giardia lamblia genome.";  
 RT Submitted (MAR-2003) to the EMBL/Genbank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AACB0100166; EAA36969.1; -; Genomic DNA.  
 SQ SEQUENCE 163 AA; 17478 MW; 3AF4A5F8DA9A10C CRC64;

Query Match 75.0%; Score 18; DB 2; Length 163;  
 Best Local Similarity 25.0%; Pred. No. 2.8e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
 Db 24 CAATTAAC 31

RESULT 12  
 Q8LDT9 ARATH PRELIMINARY; PRT; 177 AA.  
 ID Q8LDT9;  
 AC Q8LDT9;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein.  
 OS Arabidopsis thaliana (Mouse-ear cress);  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 CC Eusteroideae; Brassicales; Brassicaceae; Arabidopsis.  
 CC NCB1\_TaxID=3702;  
 OK NCB1\_TaxID=3702;  
 RN NUCLEOTIDE SEQUENCE.  
 RP MEDLINE=22086475; PubMed=12093376;  
 RX Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,  
 RA Feldmann K.A., Flavel R.B., White O., Salzberg S.L.,  
 RT "Full-length messenger RNA sequences greatly improve genome  
 RT annotation.";  
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).  
 RN NUCLEOTIDE SEQUENCE.  
 RP Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavel R.,  
 RA Feldmann K.,  
 RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AY085804; AAM63020.1; -; mRNA.  
 KM Hypothetical protein.  
 SQ SEQUENCE 177 AA; 19819 MW; CCA7841CBAE8B38 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 177;  
 Best Local Similarity 25.0%; Pred. No. 2.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
SQ CSSSSSSC 94

## RESULT 13

O9C9L7\_ARATH PRELIMINARY; PRT; 177 AA.  
AC O9C9L7\_2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Hypothetical protein T6C23.4.  
GN Name=T6C23.4;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,  
RA Maiti R., Rongling C.M., Koo H., Fujii C.Y., Uteirack T.R.,  
RA Bamsstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Town C.D., Kaul S.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC013289; AAG52545.1; -; Genomic\_DNA.  
DR PIR; F96719; F96719.  
KW Hypothetical protein.  
SQ SEQUENCE 177 AA; 19834 MW; 03D870608FEB8246 CRC64;  
Query Match 75.0%; Score 18; DB 2; Length 177;  
Best Local Similarity 25.0%; Pred. No. 2.9e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Oy 1 CXXXXXXC 8  
Db 87 CSSSSSSC 94

## RESULT 14

O60M47\_CABR PRELIMINARY; PRT; 199 AA.  
AC O60M47;  
DT 25-OCT-2004 (TREMBlrel. 28, Created)  
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
DE Hypothetical protein CBG23307.  
GN Name=CBG23307;  
OS Caenorhabditis briggsae.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodidae; Caenorhabditis.  
OX NCBI\_TaxID=6238;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA The C.briggsae Sequencing Consortium;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Contains 2 BPTI/Kunitz inhibitor domains.  
DR EMBL; CAAC01000141; CAE75329.1; -; Genomic\_DNA.  
DR GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.  
DR InterPro; IPR002223; Prot\_inh\_kunz-m.  
DR InterPro; IPR006150; Worm\_repeat\_1.  
DR Pfam; PF00014; Kunitz\_BPTI; 2.  
DR PRINTS; PR00759; BASICPRASE.  
DR PRODOM; PD000222; Prot\_inh\_kunz-m; 2.  
DR SMART; SM00131; KU; 2.  
DR SMART; SM00289; WRI; 1.  
DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 1.  
DR PROSITE; PS50279; BPTI\_KUNITZ\_2; 2.

KW Hypothetical protein.  
SQ SEQUENCE 199 AA; 21668 MW; 380A14B90A152A3 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 199;  
Best Local Similarity 25.0%; Pred. No. 3.1e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
Db 112 CSSAATTC 119

## RESULT 15

OSVMP6\_ORYSA PRELIMINARY; PRT; 200 AA.  
AC OSVMP6\_2005 (TREMBlrel. 29, Created)  
DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)  
DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)  
DE Hypothetical protein OSJNB0008D07.38.  
GN Name=OSJNB0008D07.38;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
RA Hishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,  
RA Karasawa M., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
RA Machita K., Maehara T., Mizuno H., Mizudayashi T., Mukai Y.,  
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
RA Yano M., Jiang J., Gojobori T.;  
RL "The genome sequence and structure of rice chromosome 1.";  
RL Nature 420:312-316(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsumoto T., Katayose Y.;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP006237; BAD69279.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 200 AA; 23098 MW; 0E4249F62A3678AA CRC64;  
Query Match 75.0%; Score 18; DB 2; Length 200;  
Best Local Similarity 25.0%; Pred. No. 3.1e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Oy 1 CXXXXXXC 8  
Db 102 CTAATATC 109

## RESULT 16

O9EOL7\_MOUSE PRELIMINARY; PRT; 224 AA.  
AC O9EOL7\_2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE TIS1D insertion variant (Fragment).  
GN Name=zfp3612; Synonyms=Btf2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

```

OC Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BLKS/J;
RA Cho K., Hobson K., Greenhalgh D.G.;
RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF205223; AAG45251.1; -; mRNA.
DR HSSP; P22893; IM90.
DR SMR; Q9EQL7; 1-49.
DR MGI; MGI:107945; Zfp3612.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR000571; Znf_CCH.
DR Pfam; PF00642; Zf_CCH; 1.
DR SMART; SM00356; Znf_CCH; 1.
FT NON_TER 1
FT NON_TER 224
SQ SEQUENCE 224 AA; 23018 MW; 26E09C8465A5A61E CRC64;

Query Match 75.0%; Score 18; DB 2; Length 224;
Best Local Similarity 25.0%; Pred. No. 3.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 129 CSSSASC 136

RESULT 17
Q6KRG5 ORYSA PRELIMINARY; PRT; 233 AA.
ID Q6KRG5;
AC Q6KRG5;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Glucanase-like.
GN Name=OSJNBa0052M16.18;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC
clone:OSJNBa0052M16."
RL Submitted (OCT-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP005841; BAD23651.1; -; Genomic_DNA.
DR Gramene; Q6KRG5; -; 24907 MW; E4805DB35FD0235 CRC64;
SQ SEQUENCE 233 AA; 24907 MW; E4805DB35FD0235 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 233;
Best Local Similarity 25.0%; Pred. No. 3.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 199 CAAATTAAC 206

RESULT 18
Q51XD3 MAGGR PRELIMINARY; PRT; 242 AA.
ID Q51XD3;
AC Q51XD3;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Predicted protein.
ORFNames=Mg07952.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.

```

```

OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantseang P., Baldwin J., Barry A.,
RA Bayul T., Bilshteyn B., Bloom T., Blye J., Boguslavsky L.,
RA Borowaky M., Boukhgaltier B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshtsang Y., Citroen M.,
RA Collamore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffley N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Geatin G., Gnerre S.,
RA Gurke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Huby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamysessels M., Karlsson E.,
RA Kells C., Kieu A., Kistner P., Kodira C., Kulbokas E., Labutli K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokysang T., Lokysang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Menus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelson T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunhkhong P., Pigant B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman J., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Strange-thomann N., Stavrakopoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Testaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Tsamla T., Teomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee B., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea."
RL Submitted (OCT-2003) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/Genbank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01000977; EAA53675.1; -; Genomic DNA.
SQ SEQUENCE 242 AA; 25532 MW; 3393ABEB816D10BA3 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 242;
Best Local Similarity 25.0%; Pred. No. 3.3e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8
Db 62 CAASSSC 69

RESULT 19
Q6BCP3 HPBV0 PRELIMINARY; PRT; 245 AA.
ID Q6BCP3;
AC Q6BCP3;

```

DT 25-OCT-2004 (TReMBLrel. 28, Created)  
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)  
DE Pre-S/S protein (Fragment).  
OS Hepatitis B virus.  
OC Viruses; Retro-transcribing viruses; Hepadnaviridae;  
OC Orthohepadnavirus.  
NCBI\_TaxID=10407;  
GN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Nguyen H.C., Dong S.H., Ho T.T.T., Ho H.T.D.;  
RT "Sequence analysis of pre-S/S gene segment from some clinical  
RT hepatitis B virus isolates."  
RL Submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AY665569; AAT77983.1; -; Genomic\_DNA.  
DR GO: GO:0003677; F:DNA binding; IEA.  
DR GO: GO:0003887; F:DNA-directed DNA polymerase activity; IEA.  
DR GO: GO:0006260; P:DNA replication; IEA.  
DR GO: GO:0016032; P:viral life cycle; IEA.  
DR InterPro: IPR000201; DNAPol viral N.  
DR InterPro: IPR000349; Hepvir\_surfacg.1.  
DR PANTHER: PTHR10832; Hepvir\_surfacg.1.  
DR Pfam: PF00242; DNA\_pol\_viral\_N; 1.  
DR Pfam: PF00695; vmsa; 1.  
KW Antigen.  
FT NON\_TER  
FT NON\_TER  
SQ SEQUENCE 245 AA; 26093 MW; C275D837A552298E CRC64;

Query Match 75.0%; Score 18; DB 2; Length 245;  
Best Local Similarity 25.0%; Pred. No. 3.3e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
DB 80 CASSTSSC 87

RESULT 20  
O9Y0E9 DROME PRELIMINARY; PRT; 258 AA.  
AC O9Y0E9;  
DT 01-NOV-1999 (TReMBLrel. 12, Created)  
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Salivary gland secretion protein (Fragment).  
GN Name=Sgs1; Synonyms=Sgs-1; ORNames=CG3047;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
NCBI\_TaxID=7227;  
OX [1]  
RN NUCLEOTIDE SEQUENCE.  
RA Roch G.E., Watlter S., Bornschein H., Lehmann M., Korge G.;  
RT "Structure and regulation of the salivary gland secretion protein gene  
RT Sgs-1 of Drosophila melanogaster."  
RL Genetics 0:0-0(1999).  
DR EMBL: AF156227; AAD43809.1; -; Genomic\_DNA.  
DR FlyBase: FBgn0003372; CG3047.  
DR FlyBase: FBgn0003372; Sgs1.  
FT NON\_TER  
FT NON\_TER  
SQ SEQUENCE 258 AA; 27832 MW; 0D9BF9218E6742EC CRC64;

Query Match 75.0%; Score 18; DB 2; Length 258;  
Best Local Similarity 25.0%; Pred. No. 3.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
DB 226 CTTTSSC 233

RESULT 21  
O7PR07 ANOGA PRELIMINARY; PRT; 269 AA.  
AC O7PR07;  
DT 01-MAR-2004 (TReMBLrel. 26, Created)  
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
DE ENSANGP0000001657 (Fragment).  
GN ORFNames=ENSANG0000001387;  
OS Anopheles gambiae str. PEST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;  
OC Anophelinae; Anopheles.  
NCBI\_TaxID=180454;  
GN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA STRAIN=PEST;  
RG The Anopheles gambiae Sequence Committee;  
RT "Anopheles gambiae re-annotation."  
RL Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA STRAIN=PEST;  
RG The Anopheles gambiae Sequence Committee;  
RL Submitted (APR-2004) to the EMBL/Genbank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL: AAB01008847; EAA06779.3; -; Genomic\_DNA.  
FT NON\_TER  
FT NON\_TER  
SQ SEQUENCE 269 AA; 24228 MW; 25B8FF71FD71F1F2 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 269;  
Best Local Similarity 25.0%; Pred. No. 3.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CXXXXXXC 8  
DB 55 CSSTSSC 62

RESULT 22  
O7Y198 ORYSA PRELIMINARY; PRT; 319 AA.  
AC O7Y198;  
DT 01-OCT-2003 (TReMBLrel. 25, Created)  
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)  
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)  
DE Hypothetical protein OSJNB0039018.17 (Hypothetical protein  
DE OSJNB0036M02.4).  
GN Name=OSJNB0039018.17; Synonyms=OSJNB0036M02.4;  
OS Oryza sativa (japonica cultivar-group);  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
NCBI\_TaxID=39947;  
OX [1]  
RN NUCLEOTIDE SEQUENCE.  
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,  
RA Overton II L.L., Teltzin T., Kim M.M., Bera J.J., Jin S.S.,  
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,  
RA Vanaken S.S., Riedmiller S.B., Uteckback T.T., Feldblyum T.V.,  
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,  
RA White O., Salzberg S.L., Fraser C.M.;  
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Buell R.;  
RL Submitted (MAY-2003) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Wang A., Maiti R., Jones K.M.,



RA Tallon L.J., Feldblum T.V., Tsitirin T., Bera J.J., Kim M.M., Jin S.,  
 RA Fadrosch D., Vuong H., Overton II L.L., Reardon M., Weaver B.,  
 RA Johri S., Uterback T.R., Pai G., Smith S., Mortman J., Haas B.J.,  
 RA Zhu W., Yang Q., Koo H., Zismann V., Hsiao J., Quackenbush J.,  
 RA White O., Salzberg S.L., Fraser C.M.,  
 RT "Oryza sativa chromosome 3 BAC OSJNB0036M02 genomic sequence."  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Buell R.;  
 RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC13930; AAP4639.1; -; Genomic DNA.  
 DR EMBL; AC145388; AAP89142.1; -; Genomic DNA.  
 DR Gramene; Q7Y198; -;  
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0016567; P:protein ubiquitination; IEA.  
 DR InterPro; IPR001841; Zn\_finger.  
 DR SMART; SMO0184; RING; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 319 AA; 33316 MW; FD51B063E2BAEB4 CRC64;  
 Query Match 75.0%; Score 18; DB 2; Length 319;  
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 1 CXXXXXXC 8  
 DB 296 CAAAAAAC 303  
 RESULT 23  
 ID 061A05\_CABER PRELIMINARY; PRT; 320 AA.  
 AC 061A05;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein CBG13750.  
 GN Name=CBG13750;  
 OS Caenorhabditis briggsae.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6238;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RG The C.briggsae Sequencing Consortium;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL; CAAC01000066; CABE8107.1; -; Genomic DNA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0004930; F:G-protein coupled receptor activity; IEA.  
 DR InterPro; IPR000168; Nm7TM\_chemrecept.  
 KW Hypothetical protein.  
 SQ SEQUENCE 320 AA; 36037 MW; 5E3389F6C7CF3C7B CRC64;  
 Query Match 75.0%; Score 18; DB 2; Length 320;  
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 1 CXXXXXXC 8  
 DB 302 CASSTSTC 309  
 RESULT 24  
 ID 062NM3\_HUMAN PRELIMINARY; PRT; 328 AA.  
 AC 062NM3;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein FLJ29006.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteria; Euarchontoglires; Primates; Catarrhini; Homnidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Stomach mucosa;  
 RA Nimmiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,  
 RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,  
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Suzuki Y., Hata H.,  
 RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,  
 RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,  
 RA Nagai K., Isegai T., Sugano S.,  
 RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK131040; BAC85481.1; -; mRNA.  
 SQ SEQUENCE 328 AA; 35381 MW; 7D04FC08E3270F8D CRC64;  
 Query Match 75.0%; Score 18; DB 2; Length 328;  
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 1 CXXXXXXC 8  
 DB 4 CTSASSSC 11  
 RESULT 25  
 ID 062A86\_ORYSA PRELIMINARY; PRT; 332 AA.  
 AC 062A86;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Limonene cyclase like protein.  
 GN Name=OSJNB0084107.20-2; Synonyms=OSJNB0002109.8-2;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Saeki T., Matsunoto T., Katayose Y.,  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone.OSJNB0084107.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Saeki T., Matsunoto T., Katayose Y.,  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone.OSJNB0002109.";  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP005179; BAC83914.1; -; Genomic DNA.  
 DR EMBL; AP005877; BAC81847.1; -; Genomic DNA.  
 DR Gramene; Q62456; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR004263; Exostosin.  
 DR Pfam; PF03016; Exostosin; 1.  
 SQ SEQUENCE 332 AA; 34451 MW; 57CDB6BFCC0CC5D68 CRC64;  
 Query Match 75.0%; Score 18; DB 2; Length 332;  
 Best Local Similarity 25.0%; Pred. No. 3.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 OY 1 CXXXXXXC 8  
 DB 27 CAAAAAAC 34

```

RESULT 26
094589 9CILI PRELIMINARY; PRT; 350 AA.
ID 094589 9CILI PRELIMINARY; PRT; 350 AA.
AC 094589
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Lemnabion factor.
OS Lemnabion bullinum.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Pericultida;
OC Lemnabion.
NCBI_TaxID=54108;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Peters-Regehr T., Kusch J., Heckmann K.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Y09220; CAAT70420.1; -, genomic_DNA.
DR HSSP; O16119; IEZG.
DR InterPro; IPR006212; Furin_repeat.
DR SMART; SM00261; FU; 3.
SQ SEQUENCE 350 AA; 35159 MW; DBR0C67654BD9D2E CRC64;

Query Match 75.0%; Score 18; DB 2; Length 350;
Best Local Similarity 25.0%; Pred. No. 3.8e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8
Db 105 CSTGATTC 112

RESULT 27
RHOM DROME STANDARD; PRT; 355 AA.
ID RHOM DROME STANDARD; PRT; 355 AA.
AC P20350; Q9W0F2;
DT 01-FEB-1991 (Rel. 17, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Rhomboid protein (EC 3.4.21.-) (Veinlet protein).
GN Name=rho; Synonyms=Ve; ORFNames=CG1004;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=90249726; PubMed=2110920;
RA Bier E., Jan L.Y., Jan Y.N.;
RT "Rhomboid, a gene required for dorsoventral axis establishment and
peripheral nervous system development in Drosophila melanogaster.";
RL Genes Dev. 4:190-203(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STAIN=Beckley;
RA MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Berens P.V., Bernier B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
Buttis K.C., Busam D.A., Butler H., Cadieu R.G., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablo B., DeJcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Dudin K.J., Evangelista C.C., Ferraz C., Petriera S., Fleischmann W.,
Fowler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasner K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

```

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jajali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Kerchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter U., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [4]
RP FUNCTION, AND MUTAGENESIS OF TRP-151; ARG-152; ASN-169; GLY-215;
RP SER-217 AND HIS-281.
RX MEDLINE=21526629; PubMed=11672525; DOI=10.1016/S0092-8674(01)00525-6;
RA Urban S., Lee J.R., Freeman M.;
RT "Drosophila Rhomboid-1 defines a family of putative intramembrane
serine proteases.";
RL Cell 107:173-182(2001).
CC -1- FUNCTION: Acts early in embryonic development to establish
position along the dorsoventral axis and then again later to
specify the fate of neuronal precursor cells. Involved in EGF
receptor signaling; cleaves Spitz to release the active growth
factor.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Golgi.
CC -1- DEVELOPMENTAL STAGE: Early blastoderm stages and later during
nervous development.
CC -1- SIMILARITY: Belongs to the peptidase S54 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; X52454, CA36692.1; -, mRNA.
CC EMBL; AE003471, AA47496.1; -, genomic_DNA.
CC PIR; A34597, A34597.
CC MEROPS; S54.001; -.
CC Ensembl; CG1004; Drosophila melanogaster.
CC FlyBase; FBgn0004635; rho.
CC GO; GO:0005794; C:Golgi apparatus; IDA.
CC GO; GO:0016021; C:Integral to membrane; TAS.
CC GO; GO:0008236; F:serine-type peptidase activity; IDA.
CC GO; GO:0007420; P:brain development; IMP.
CC GO; GO:0046845; P:branched duct epithelial cell fate determin. . .; TAS.
CC GO; GO:0001763; P:branching morphogenesis; NAS.
CC GO; GO:0007143; P:epidermal growth factor ligand processing; IGI.
CC GO; GO:0007479; P:leg disc proximal/distal pattern formation; TAS.
CC GO; GO:0007438; P:neocyte development; IGI.

```

DR GO: GO:0030707; P:ovarian follicle cell development (sensu In. . .; TAS.  
 DR GO: GO:0007442; P:peripheral nervous system development; TAS.  
 DR GO: GO:0045742; P:positive regulation of epidermal growth fac. . .; TAS.  
 DR GO: GO:0006508; P:proteolysis and peptidolysis; ICI.  
 DR GO: GO:0007166; P:regulation of epidermal growth factor recep. . .; TAS.  
 DR GO: GO:0007432; P:salivary gland determination; NAS.  
 DR GO: GO:0005202; P:tracheal sac formation (sensu Insecta); TAS.  
 DR GO: GO:0007474; P:wing vein specification; NAS.  
 DR InterPro: IPR002610; Rhomboid-like.  
 DR Pfam: PF01694; Rhomboid.1.  
 KM Developmental protein; Golgi stack; Hydrolase; Protease;  
 KW Serine protease; Transmembrane.  
 FT TOPO\_DOM 1 98 Cytoplasmic (Potential).  
 FT TRANSMEM 99 119 Potential.  
 FT TOPO\_DOM 120 162 Luminal (Potential).  
 FT TRANSMEM 163 183 Potential.  
 FT TOPO\_DOM 184 188 Cytoplasmic (Potential).  
 FT TRANSMEM 189 209 Potential.  
 FT TOPO\_DOM 210 210 Luminal (Potential).  
 FT TRANSMEM 211 231 Potential.  
 FT TOPO\_DOM 232 244 Cytoplasmic (Potential).  
 FT TRANSMEM 245 265 Potential.  
 FT TOPO\_DOM 266 275 Luminal (Potential).  
 FT TRANSMEM 276 296 Potential.  
 FT TOPO\_DOM 297 308 Cytoplasmic (Potential).  
 FT TRANSMEM 309 329 Potential.  
 FT TOPO\_DOM 330 355 Luminal (Potential).  
 FT ACT\_SITE 169 169 Charge relay system.  
 FT ACT\_SITE 217 217 Charge relay system.  
 FT ACT\_SITE 281 281 Charge relay system.  
 FT MUTAGEN 151 151 W->A: Abolishes protease activity.  
 FT MUTAGEN 152 152 R->A: Abolishes protease activity.  
 FT MUTAGEN 169 169 N->A: Abolishes protease activity.  
 FT MUTAGEN 215 215 G->A: Abolishes protease activity.  
 FT MUTAGEN 217 217 S->C,T: Abolishes protease activity.  
 FT MUTAGEN 281 281 H->A: Abolishes protease activity.  
 FT CONFLICT 4 4 L -> P (in Ref. 1).  
 FT CONFLICT 33 33 A -> V (in Ref. 1).  
 FT CONFLICT 46 46 S -> T (in Ref. 1).  
 SQ SEQUENCE 355 AA; 39330 MW; 2E0572D08831C5D0 CRC64;

Query Match 75.0%; Score 18; DB 1; Length 355;  
 Best Local Similarity 25.0%; Pred. No. 3.8e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
 DB 57 CSTASSTC 64

RESULT 28  
 Q540V7 DROME PRELIMINARY; PRT; 355 AA.  
 AC Q540V7;  
 DT 13-SEP-2005 (TREMBLrel. 31, Created)  
 DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)  
 DE LDD6131P.  
 GN Name=rho;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 OK NCB1\_Taxid=7227;  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=Berkeley.  
 RC Striplaton M., Brockstein P., Hong L., Aghayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese B.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liso G.,  
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Patagas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceiniker S.;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY121641; AAM51968.1; -, mRNA.  
 SQ SEQUENCE 355 AA; 39330 MW; 2E0572D08831C5D0 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 355;  
 Best Local Similarity 25.0%; Pred. No. 3.8e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8  
 DB 57 CSTASSTC 64

RESULT 29  
 T1SD MOUSE STANDARD; PRT; 367 AA.  
 AC P23949;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Butyrate response factor 2 (T1SID protein).  
 GN Name-zfp3612; Synonyms=Brf2, Tis1d;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=BALB/c;  
 RX MEDLINE=91141531; PubMed=1996120;  
 RA Varnum B.C., Ma Q., Chi T., Fletcher B., Herschman H.R.;  
 RT "The T1SID primary response gene is a member of a gene family that  
 RT encodes proteins with a highly conserved sequence containing an  
 RT unusual Cys-His repeat.";  
 RL Mol. Cell. Biol. 11:1754-1758(1991).  
 RN [2]  
 NP NUCLEOTIDE SEQUENCE OF 1-50.  
 RA Fletcher B.S.;  
 RL Thesis (1992); University of California Los Angeles, United States.  
 CC -1- FUNCTION: Probable regulatory protein involved in regulating the  
 CC response to growth factors.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -1- SIMILARITY: Contains 2 C3H1-type zinc fingers.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 DR EMBL; M58564; AAA72946.1; -, mRNA.  
 DR EMBL; M97165; AAA39709.1; -, Genomic\_DNA.  
 DR PIR; C39590; C39590.  
 DR HSSP; P22893; 1M90.  
 DR SMK; P23949; 124-193.  
 DR Ensembl; ENSMUSG00000045817; Mus musculus.  
 DR MGI; MGI:107945; Zfp3612.  
 DR InterPro; IPR007635; T1SID N.  
 DR InterPro; IPR00571; Znf\_CCH.  
 DR Pfam; PF04553; T1SID C/1.  
 DR Pfam; PF00642; zf-CCH1; 2.  
 DR SMART; SM00356; Znf\_C3H1; 2.  
 DR DNA-binding; Metal-Binding; Nuclear protein; Repeat; Zinc;  
 KW Zinc-finger.  
 FT ZN\_FING 132 151 C3H1-type 1.  
 FT ZN\_FING 170 189 C3H1-type 2.  
 FT COMPBIAS 64 67 Poly-Gly.  
 FT COMPBIAS 111 114 Poly-Gln.  
 FT COMPBIAS 200 203 Poly-Gly.  
 FT COMPBIAS 263 266 Poly-Pro.  
 FT COMPBIAS 297 303 Poly-Ala.  
 SQ SEQUENCE 367 AA; 37593 MW; 361244AF6244E46E CRC64;

Query Match 75.0%; Score 18; DB 1; Length 367;  
 Best Local Similarity 25.0%; Pred. No. 3.9e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
 Db 273 C5SSASSC 280

## RESULT 30

086SB6\_DROVI PRELIMINARY; PRT; 404 AA.  
 AC 086SB6;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Rhomboid.  
 GN Name=rho;  
 OS Drosophila virilis (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7244;  
 RX MEDLINE=22844358; PubMed=12963110; DOI=10.1016/S0925-4773(03)00164-3;  
 RA Nakamura Y., Matsuno K.;  
 RT "Species-specific activation of EGF receptor signaling underlies  
 RT evolutionary diversity in the dorsal appendage number of the genus  
 RT Drosophila egsheila.";  
 RL Mech. Dev. 120:897-907(2003).  
 DR EMBL; AB089248; BACS6701.1; -; mRNA.  
 DR MEROPS; S54.001; -;  
 DR FlyBase; FBgn0062278; DvYrhc.  
 DR InterPro; IPR002610; Rhomboid-like.  
 DR Pfam; PF01694; Rhomboid; 1.  
 SQ SEQUENCE 404 AA; 44035 MW; EF641632B02F9011 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 404;  
 Best Local Similarity 25.0%; Pred. No. 4e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
 Db 63 C5TASSTC 70

RESULT 31  
 06ZAE5\_ORYSA PRELIMINARY; PRT; 417 AA.  
 AC 06ZAE5;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Limonene cyclase like protein.  
 GN Name=OSJNB0084L07.20-3; Synonyms=OSJNB0002L09.8-3;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Erihartoideae; Oryzaceae; Oryza.  
 NCBI\_TaxID=39947;  
 RX MEDLINE=39947;  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RT "Oryza sativa nipponbata (G43) genomic DNA, chromosome 7, BAC  
 RT clone:OSJNB0084L07.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RT "Oryza sativa nipponbata (G43) genomic DNA, chromosome 7, BAC  
 RT clone:OSJNB0002L09.";  
 OS Name=OSJNB0002L09.";

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP005179; BAC83915.1; -; Genomic DNA.  
 DR EMBL; AP005877; BAC31849.1; -; Genomic DNA.  
 DR Gramene; 06ZAE5; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR004263; Exostosin.  
 DR Pfam; PF03016; Exostosin; 1.  
 SQ SEQUENCE 417 AA; 44610 MW; 67668E9A53BIDC6A CRC64;

Query Match 75.0%; Score 18; DB 2; Length 417;  
 Best Local Similarity 25.0%; Pred. No. 4.1e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
 Db 27 CAAAAC 34

## RESULT 32

08UTL4\_ORYLA PRELIMINARY; PRT; 445 AA.  
 ID 08UTL4\_ORYLA  
 AC 08UTL4;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE TAPBP Protein.  
 GN Name=TAPBP;  
 OS Oryzias latipes (Medaka fish) (Japanese ricefish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
 OC Belontiiformes; Adrianchthyidae; Oryziinae; Oryzias.  
 NCBI\_TaxID=8090;  
 RX MEDLINE=21850510; PubMed=11862394; DOI=10.1007/s00251-001-0427-3;  
 RA Matsuo M.Y., Asakawa S., Shimizu N., Kimura H., Nonaka M.;  
 RT "Nucleotide sequence of the MHC class I genomic region of a teleost,  
 RT the medaka (Oryzias latipes).";  
 RL Immunogenetics 53:930-940(2002).  
 DR EMBL; BA000027; BAB83851.1; -; Genomic DNA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0019885; P:antigen processing, endogenous antigen via . . .; IEA.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR008056; Tapasin.  
 DR Pfam; PF00047; IG; 1.  
 DR PRINTS; PRO1669; TAPASIN.  
 DR SMART; SM00409; IG; 2.  
 DR PROSITE; PS50835; IG LIKE; 2.  
 KW Immunoglobulin domain; Repeat.  
 SQ SEQUENCE 445 AA; 48260 MW; FDF7C5FBE74370BB CRC64;

Query Match 75.0%; Score 18; DB 2; Length 445;  
 Best Local Similarity 25.0%; Pred. No. 4.2e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 8  
 Db 22 C5SSSSSC 29

## RESULT 33

07SHC6\_NEUCR PRELIMINARY; PRT; 480 AA.  
 ID 07SHC6\_NEUCR  
 AC 07SHC6;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Predicted protein (Hypothetical protein B13N4.210).  
 GN Name=NCU01880.1; Synonyms=B13N4.210;  
 OS Neurospora crassa.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OC NCBI\_TaxID=5141;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=OR74A;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Puccelli S., Rehman B.,  
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Janakiev P., Pedersen D., Nelson M., Washburne M.,  
 RA Seltrennkoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
 RA Koche G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
 RA Kamal M., Kamysseilis M., Mauceli E., Bielke C., Rudd S., Frishman D.,  
 RA Kyrstova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken D.,  
 RA Cogoni C., Marino G., Catchside D., Li W., Pratt R.U., Osmann S.A.,  
 RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
 RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,  
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa."  
 RL Nature 0:0-0(2003).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Schulte U., Aign V., Hohesl J., Brandt P., Fartmann B., Holland R.,  
 RA Nykatura G., Mewes H.W., Mannhaupt G.,  
 RL Submitted (DEC-2003) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RA German Neurospora genome project;  
 RL Submitted (DEC-2003) to the EMBL/Genbank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AABX01000004; EAA36275.1; -; Genomic\_DNA.  
 DR EMBL; BX842681; CAE81980.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 480 AA; 50449 MW; E073DB8EC33C9A2B CRC64;  
 Qy 1 CXXXXXXC 8  
 Db 127 CSSSSSSC 134  
 Query Match 75.0%; Score 18; DB 2; Length 480;  
 Best Local Similarity 25.0%; Pred. No. 4.3e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 RESULT 34  
 ID 0758Y3 ASHGO PRELIMINARY; PRT; 480 AA.  
 AC 0758Y3;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE AD3395CD.  
 GN Name=ADR395C;  
 OS Ashbya gossypii (Yeast) (Eremothecium gossypii).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Eremothecium.  
 OC NCBI\_TaxID=33169;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=ATCC 10895;  
 RX PubMed=15001715; DOI=10.1126/science.1095781;  
 RA Dietrich F.S., Voegel S., Brachat S., Lerch A., Gates K., Steiner S.,  
 RA Morf C., Poehlmann R., Iuedi P., Choi S., Wing R.A., Flavier A.,  
 RA Gaffney T.D., Philippson P.,  
 RT "The Ashbya gossypii genome as a tool for mapping the ancient  
 RT Saccharomycetes ceterisibae genome."  
 RL Science 304:304-307(2004).  
 DR EMBL; AE016817; AAS52314.1; -; Genomic\_DNA.  
 DR ACD; ADR395C; -.

DR GO; GO:000414; F:3', 5'-cyclic-nucleotide phosphodiesterase a. .; IEA.  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR003607; Met\_phos\_hydro.  
 DR InterPro; IPR02073; PDBase.  
 DR Pfam; PF00233; PDBase\_1; 1.  
 DR PRINTS; PR00387; PDIESTERASE1.  
 DR SMART; SM00471; HDc.1.  
 DR PROSITE; PS00126; PDBASE\_1; UNKNOWN\_1.  
 KW COMPLETE PROTEOME.  
 SQ SEQUENCE 480 AA; 53835 MW; 75AB3894A55D01CB CRC64;  
 Qy 1 CXXXXXXC 8  
 Db 205 CTAATAC 212  
 Query Match 75.0%; Score 18; DB 2; Length 480;  
 Best Local Similarity 25.0%; Pred. No. 4.3e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 RESULT 35  
 ID T1SD HUMAN STANDARD; PRT; 492 AA.  
 AC P47974; Q9BSJ3;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Butyrate response factor 2 (T1SID protein) (BGF-response factor 2)  
 DE (BRF-2).  
 GN Name=ZFP36L2; Synonyms=BRF2, BRF2, T1SID;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=96137224; PubMed=8545129;  
 RA Ino T., Yasui H., Hirano M., Kurosawa Y.,  
 RT "Identification of a member of the T1SID early response gene family at  
 RT the insertion point of a DNA fragment containing a gene for the T-cell  
 RT receptor beta chain in an acute T-cell leukemia."  
 RL Oncogene 11:2705-2710(1995).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=95137407; PubMed=7835719; DOI=10.1016/0378-1119(94)00636-P;  
 RA Nie X.F., Maclean K.N., Kumar V., McKay I.A., Bustin S.A.;  
 RT "BRF-2, the human homologue of the murine T1sid early response  
 RT gene".  
 RL Gene 152:285-286(1995).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,  
 RA Alechuth S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueidi T.B., Toshitsuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Matek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green B.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Merra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences." ;  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- FUNCTION: Probable regulatory protein involved in regulating the  
CC response to growth factors.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -1- SIMILARITY: Contains 2 C3H1-type zinc fingers.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL; X7892; CAAS5592.1; -; mRNA.  
DR EMBL; BC005010; AAH05010.1; -; mRNA.  
DR PIR; S49147; S49147.  
DR PDB; 1RGO; NMR; A=151-220.  
DR Ensembl; ENSG00000152518; Homo sapiens.  
DR HGNC; HGNC:1108; ZFP36L2.  
DR H-InVDB; HIX0002007; -.  
DR GO; GO:0003700; F:transcription factor activity; TAS.  
DR GO; GO:0008283; P:cell proliferation; TAS.  
DR InterPro; IPR007635; Tis11B\_N.  
DR InterPro; IPR005711; Znf\_CCH.  
DR Pfam; PF04553; Tis11B\_C; 1.  
DR Pfam; PF00642; zf-CCCH; 2.  
DR SMART; SM00356; Znf\_C3H1; 2.  
KW 3D-structure; DNA-binding; Metal-binding; Nuclear protein; Repeat;  
KW Zinc; Zinc-finger.  
FT ZN\_FING 159 178 C3H1-type 1.  
FT ZN\_FING 197 216 C3H1-type 2.  
FT COMPBIAS 106 109 Poly-Gln.  
FT COMPBIAS 138 141 Poly-Gln.  
FT COMPBIAS 143 146 Poly-Gly.  
FT COMPBIAS 288 291 Poly-Pyro.  
FT COMPBIAS 323 330 Poly-Ala.  
FT COMPBIAS 382 388 Poly-Ala.  
FT COMPBIAS 393 399 Poly-Gln.  
FT CONFLICT 96 97 TS -> DL (in Ref. 1).  
FT CONFLICT 318 318 AA -> T (in Ref. 2).  
FT CONFLICT 329 330 AA -> LR (in Ref. 2).  
FT CONFLICT 330 330 A -> AAA (in Ref. 3).  
FT CONFLICT 399 399 Q -> QQQQ (in Ref. 3).  
FT CONFLICT 451 460 Missing (in Ref. 1).  
SQ SEQUENCE 492 AA; 50921 MW; E4E2EE266791CCF7 CRC64;  
Query Match 75.0%; Score 18; DB 1; Length 492;  
Best Local Similarity 25.0%; Pred. No. 4.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 CXXXXXXC 8  
Db 299 CSSASASC 306

RESULT 36  
053TB4\_HUMAN PRELIMINARY; PRT; 494 AA.  
AC 053TB4;  
DT 13-SEP-2005 (TREMBlrel. 31, Created)  
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)  
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)  
DE Hypothetical protein ZFP36L2.  
GN Name=ZFP36L2;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Abbott A., Boyer E., Heyen J.;

RT "The sequence of Homo sapiens BAC clone RP11-339H12." ;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Waterston R.H.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RA Waterston R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RA Wilson R.K.;  
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC010883; AA14992.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 494 AA; 51062 MW; 10E23FA9E2DDABD4 CRC64;  
Query Match 75.0%; Score 18; DB 2; Length 494;  
Best Local Similarity 25.0%; Pred. No. 4.4e+02;  
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 CXXXXXXC 8  
Db 299 CSSASASC 306

RESULT 37  
06PD21\_MOUSE PRELIMINARY; PRT; 526 AA.  
AC 06PD21;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Shp protein (Fragment).  
GN Name=Shp;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
STRAIN=C57BL/6; TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derse J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heileh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Tohiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnaracne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Huliy S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalek U., Smallin D.E.,  
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences." ;  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
STRAIN=C57BL/6; TISSUE=Brain;  
RA Strauberg R.;  
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC058986; AAH58986.1; -; mRNA.  
DR HSSP; P00524; IKC2.  
DR GO; GO:0005515; F:protein binding; IDA.

```

DR GO:0005070; F:SH3/SH2 adaptor activity; TAS.
DR GO:0042100; P:B-cell proliferation; IDA.
DR InterPro: IPR00980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRODOM: PD00093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS0001; SH2; 1.
FT NON TER
SQ SEQUENCE 526 AA; 57393 MW; 6F358E391E2C63E0 CRC64;

Query Match
Best Local Similarity 75.0%; Score 18; DB 2; Length 526;
Pred. No. 4.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 71 CSASSASC 78

RESULT 38
OSB195_DROME PRELIMINARY; PRT; 544 AA.
ID Q5B195_DROME PRELIMINARY; PRT; 544 AA.
AC Q5B195;
DT 10-MAY-2005 (TREMBLrel. 30, Created)
DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE R604051P.
GN Name=dpa;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkley;
RA Stimpson M., Carlson J., Chavez C., Friese B., George R., Pacleb J.,
RA Pak S., Wan K., Yu C., Rubin G.M., Celinker S.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT021329; AAX33477.1; -; mRNA.
DR GO:0005634; C:nucleus; IEA.
DR GO:0005524; F:ATP binding; IEA.
DR GO:0003677; F:DNA binding; IEA.
DR GO:0008094; F:DNA replication; IEA.
DR GO:0006270; F:DNA replication initiation; IEA.
DR InterPro: IPR01304; Lectin_C.
DR InterPro: IPR01208; MCM_4.
DR InterPro: IPR008047; MCM_4.
DR PRINTS: PR01660; MCMPROTEIN4.
DR SMART: SM00350; MCM; 1.
DR PROSITE: PS00615; C-TYPE LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50051; MCM_2; 1.
SQ SEQUENCE 544 AA; 60347 MW; 8D221F998960D707 CRC64;

Query Match
Best Local Similarity 75.0%; Score 18; DB 2; Length 544;
Pred. No. 4.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 522 CSTSTSC 529

RESULT 39
Q91TR2_TUHV1 PRELIMINARY; PRT; 545 AA.
ID Q91TR2_TUHV1 PRELIMINARY; PRT; 545 AA.
AC Q91TR2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DT T31.
OS Tupaiid herpesvirus 1 (strain 1) (TuHV-1) (Herpesvirus tupaia (strain

```

```

OS 1)).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OC NCBI_TaxID=10397;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=2;
RX MEDLINE=21211637; PubMed=11312357;
RX DOI=10.1128/JVI.75.10.4854-4870.2001;
RA Bahr U., Darai G.;
RT "Analysis and characterization of the complete genome of tupaia (tree
RT shrew) herpesvirus."
RT J. Virol. 75:4854-4870(2001).
DR EMBL; AF281817; AAK57075.1; -; Genomic_DNA.
DR InterPro: IPR007578; DUF570.
DR Pfam: PF04489; DUF570; 1.
SQ SEQUENCE 545 AA; 59508 MW; 98B5EC0DEF7A806 CRC64;

Query Match
Best Local Similarity 75.0%; Score 18; DB 2; Length 545;
Pred. No. 4.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 535 CAASATC 542

RESULT 40
Q5Z8N9_ORYSA PRELIMINARY; PRT; 551 AA.
ID Q5Z8N9_ORYSA PRELIMINARY; PRT; 551 AA.
AC Q5Z8N9;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Aspartic proteinase nepenthesin II-like.
GN Name=P0541C02.19-2;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhacoidae; Oryzae; Oryza.
OC NCBI_TaxID=39947;
OX [1]
RN NUCLEOTIDE SEQUENCE.
RP Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone: P0541C02.19-2."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003769; BAD61723.1; -; Genomic_DNA.
DR GO:0004194; F:pepsin A activity; IEA.
DR GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR01461; Peptidase A1.
DR InterPro: IPR01969; Pept_Asp_AS.
DR Pfam: PF00026; Asp; 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP PROTEASE; UNKNOWN_2.
SQ SEQUENCE 551 AA; 57586 MW; BE7840EB14824DF CRC64;

Query Match
Best Local Similarity 75.0%; Score 18; DB 2; Length 551;
Pred. No. 4.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
DB 182 CATATSC 189

RESULT 41
Q8SB30_ORYSA PRELIMINARY; PRT; 551 AA.
ID Q8SB30_ORYSA PRELIMINARY; PRT; 551 AA.
AC Q8SB30;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

```

DE Putative chloroplast nucleoid DNA-binding protein.  
 GN Name=OJ1540.H01.13;  
 OS Oryza sativa (Rice).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC Euphorbiaceae; Oryzae; Oryza.  
 CC NCBI\_TaxID=4530;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RA Buell C.R., Yuan O., Ouyang S., Liu J., Moffat K.S., Hill J.N.,  
 RA Gansberger K., Bremner M., Burgess S., Hance M., Shvartsbeyn M.,  
 RA Tselirin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,  
 RA Vanakke S.E., Utterback T.R., Feldblum T.V., Kalb E., Quackenbush J.,  
 RA Salzberg S.L., White O., Fraser C.M.;  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP NUCLEOTIDE SEQUENCE.  
 RA Buell R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC091774; AL79734.1; -, Genomic\_DNA.  
 DR HSSP; P00797; 2REN.  
 DR Gramene; O8S30; -;  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004194; F:pepsin A activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR01461; Peptidase A1.  
 DR InterPro; IPR01969; Pept\_Asp\_AS.  
 DR Pfam; PF00026; Asp; 1.  
 DR PRINTS; PR00792; PEPsin.  
 DR PROSITE; PS00141; ASP\_PROTASE; UNKNOWN\_2.  
 DR DNAS-binding.  
 KW SEQUENCE 551 AA; 57586 MW; BE78840EB14824DF CRC64;  
 SQ  
 Query Match 75.0%; Score 18; DB 2; Length 551;  
 Best Local Similarity 25.0%; Pred. No. 4.6e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 CXXXXXXC 8  
 DB 182 CATATSSC 189

RESULT 42  
 Q5AP57 CANAL PRELIMINARY; PRT; 552 AA.  
 ID Q5AP57;  
 AC Q5AP57;  
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
 DE Hypothetical yfw family protein 5.  
 GN ORFNames=Ca019.4881;  
 OS Candida albicans SC5314.  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 CC NCBI\_TaxID=237561;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RA STRAIN=SC5314;  
 RX PubMed=15123810; DOI=10.1073/pnas.0401648101;  
 RA Jones T., Federpiel N.A., Chibana H., Dungan J., Kalman S.,  
 RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,  
 RA Davis R.W., Scherer S.;  
 RT "The diploid genome sequence of Candida albicans";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334 (2004).  
 RN (2)  
 RP NUCLEOTIDE SEQUENCE.  
 RA STRAIN=SC5314;  
 RA Dungan U., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegoke O.,  
 RA Roberts J., Petersen K., Donnelly S., Favoreto S., Tzung K.-W.,  
 RA Jones T., Scherer S., Agabian N.;  
 RT "Annotation of the Genome of Candida albicans";  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAC01000001; EAL04869.1; -, Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 552 AA; 62256 MW; 5CA72B0ED5B87B3F CRC64;  
 Query Match 75.0%; Score 18; DB 2; Length 552;  
 Best Local Similarity 25.0%; Pred. No. 4.6e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 CXXXXXXC 8  
 DB 466 CTSATTSC 473

RESULT 43  
 Q80J53 HPBV0 PRELIMINARY; PRT; 560 AA.  
 ID Q80J53; HPBV0  
 AC Q80J53;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Truncated polymerase.  
 OS Hepatitis B virus.  
 CC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 CC NCBI\_TaxID=10407;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RA Lin X., Xu X., Zheng D.;  
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY238972; AAC04453.1; -, Genomic\_DNA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0003887; F:DNA-directed DNA polymerase activity; IEA.  
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR GO; GO:0003964; F:RNA-directed DNA polymerase activity; IEA.  
 DR GO; GO:0016740; F:transference activity; IEA.  
 DR GO; GO:0006278; P:RNA-dependent DNA replication; IEA.  
 DR InterPro; IPR001462; DNAPol\_viral\_C.  
 DR InterPro; IPR000201; DNAPol\_viral\_N.  
 DR InterPro; IPR00477; RTase.  
 DR Pfam; PF00336; DNA\_pol\_viral\_C; 1.  
 DR Pfam; PF00242; DNA\_pol\_viral\_N; 1.  
 DR Pfam; PF00078; RVT\_1; 1.  
 SQ SEQUENCE 560 AA; 62448 MW; 7BCD0F6F891CE4B CRC64;  
 Query Match 75.0%; Score 18; DB 2; Length 560;  
 Best Local Similarity 25.0%; Pred. No. 4.6e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 CXXXXXXC 8  
 DB 268 CASSSSC 275

RESULT 44  
 Q6N1S1 CORDI PRELIMINARY; PRT; 581 AA.  
 ID Q6N1S1; CORDI  
 AC Q6N1S1;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Putative secreted protein.  
 GN OrderedclonNames=DIP0696;  
 OS Corynebacterium diphtheriae.  
 CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 CC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
 CC NCBI\_TaxID=1717;  
 RN (1)  
 RP NUCLEOTIDE SEQUENCE.  
 RA STRAIN=Biotype gravis / NCTC 13129;  
 RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;  
 RA Cerdano-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,



RA Pallen M.J., Bentley S.D., Beara G.S., Churcher C.M., James K.D.,  
 RA de Zorja A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,  
 RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,  
 RA Rabinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,  
 RA Whitehead S., Barrett B.G., Parkhill J.,  
 RT "The complete genome sequence and analysis of *Corynebacterium*  
 RT *diphtheriae* NCTC13129."  
 RL Nucleic Acids Res. 31:6516-6523 (2003).  
 DR EMBL; BX248355; CAB49213.1; -; Genomic\_DNA.  
 KW Complete proteome.  
 SQ SEQUENCE 581 AA; 62851 MW; 9CF98BC85A311AC3 CRC64;

Query Match 75.0%; Score 18; DB 2; Length 581;  
 Best Local Similarity 25.0%; Pred. NO. 4.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
 Db 17 CAATTAAC 24

RESULT 45  
 Q51VZ0\_MAGGR PRELIMINARY; PRT; 592 AA.  
 AC Q51VZ0;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=MG033995.4;  
 OS Magnaporthe grisea 70-15.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetes; incertae sedis; Magnaporthaceae; Magnaporthae.  
 NCBI\_TaxID=242507;  
 OX [1]

NUCLEOTIDE SEQUENCE.

RA STRAIN=70-15;  
 RA Alt-zahra M., Nusham C., Abebe A., Abouelleil A., Adekoya E.,  
 RA Birren B., Allen N., Allen T., An P., Anderson M., Anderson S.,  
 RA Archchi H., Ambuster J., Bachantsang P., Baldwin J., Barry A.,  
 RA Bayul T., Bilshteyn B., Bloom T., Blye J., Boguslavskiy L.,  
 RA Bowdley M., Boulghalfer B., Brunache A., Butler J., Calixte N.,  
 RA Calvo S., Camarata J., Campo K., Chang J., Cheahatsang Y., Citroen M.,  
 RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,  
 RA David R., Dawe T., Degray S., Dodge S., Dooley K., Dorje P.,  
 RA Dorjee K., Dorris L., Dufley N., Dupes A., Elkins T., Engels R.,  
 RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,  
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Geatin G., Gnerre S.,  
 RA Gutirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,  
 RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,  
 RA Homan T., Horn A., Houde N., Hughes L., Kanyseilis M., Karlsson E.,  
 RA Jaffe D., Jones C., Kamen A., Kodira C., Kulbokas E., Labutti K.,  
 RA Kells C., Kien A., Kiser P., Kodira C., Lewis D., Lewis T.,  
 RA Lama D., Landers T., Legner J., Levine S., Lewis D., Lewis T.,  
 RA Lindblad-Toh K., Liu X., Loksitsang T., Loksitsang Y., Lucien O.,  
 RA Lui A., Ma L.J., Mahitt R., Macdonald J., Maclean C., Major J.,  
 RA Manning J., Marabelli A., Maru K., Matthews C., Mauceli E.,  
 RA McCarthy M., McDonough S., McGhee T., Meldrum J., Menus L.,  
 RA Meitov J., Mihalev A., Minova T., Mikkelson T., Mlenga V., Moru K.,  
 RA Moses J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,  
 RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizari M., Norby C.,  
 RA Norbu N., O'Donnell P., Okawa O., O'Leary S., Omotohio B.,  
 RA O'Neill K., Osman S., Parker S., Perrin D., Phunhang P., Pignani B.,  
 RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,  
 RA Rector R., Richardson R., Rise C., Rodriguez J., Rogers J., Rogov P.,  
 RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,  
 RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnaz C.,  
 RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,  
 RA Stetson K., Stone C., Stone S., Strube M., Talmes J., Thunga P.,  
 RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,  
 RA Towey S., Tsamla T., Tesomo N., Vallee D., Vassiliev H.,  
 RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,  
 RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,

RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,  
 RA Zimmer A., Zody M., Lander E.,  
 RT "The genome sequence of *Magnaporthe grisea*."  
 RT Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=70-15;

RA Dean R., Mitchell T., Brown D., Pan H., Thon M.,

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=70-15;

RA Zhu H., Blackmon B.,

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

CC -! CAUTION: The sequence shown here is derived from an

EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

preliminary data.

DR EMBL; AACU01001048; EAA50236.1; -; Genomic\_DNA.

KW Hypothetical protein.

SQ SEQUENCE 592 AA; 63027 MW; 84D30AEC5A5D8ABC CRC64;

Query Match 75.0%; Score 18; DB 2; Length 592;  
 Best Local Similarity 25.0%; Pred. NO. 4.7e+02;  
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 8  
 Db 347 CAAASTGC 354

RESULT 46  
 O6Z4E7\_ORYSA PRELIMINARY; PRT; 606 AA.  
 ID O6Z4E7;  
 AC O6Z4E7;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Putative pectin-glucuronyltransferase.  
 GN Name=OSUNB0084107.20-1; Synonyms=OSUNB0084107.8-1;  
 OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriatoidae; Oryzaceae; Oryza.  
 NCBI\_TaxID=39947;  
 OX [1]

NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Katayose Y.,  
 RT "Oryza sativa nippohare (GA3) genomic DNA, chromosome 7, BAC  
 RT clone:OSUNB0084107."  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.



```

GN Name=YFW5; ORFNames=CaO19.12344;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=237561;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RT "The diploid genome sequence of Candida albicans.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RA Dungan J., Kuo A., Newport G., Ian C.-Y., Iijima C., Adegboola O.,
RA Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
RA Jones T., Scherer S., Agabian N.;
RT "Annotation of the Genome of Candida albicans.";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAC001000002; EAL04673.1; -; Genomic_DNA.
KW Hypothetical protein.
SO SEQUENCE 722 AA; 81770 MW; 16E7C88A35A416AE CRC64;

Query Match 75.0%; Score 18; DB 2; Length 722;
Best Local Similarity 25.0%; Pred. No. 5.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CXXXXXXC 8
Db 636 CTSARTSC 643

```

Search completed: January 4, 2006, 16:09:52  
 Job time : 54.2435 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 ; Search time 61.5913 Seconds  
(without alignments)  
64.204 Million cell updates/sec

Title: US-09-932-322-10

Perfect score: 25

Sequence: 1 CXXXXXXC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 1000 summaries

Database :

1: A\_Geneseq\_21.\*  
2: geneseqp19808.\*  
3: geneseqp20008.\*  
4: geneseqp20018.\*  
5: geneseqp20028.\*  
6: geneseqp20038.\*  
7: geneseqp20048.\*  
8: geneseqp20058.\*  
9: geneseqp20068.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	72.0	59	4	ABBA1147 Peptide #
2	18	72.0	59	4	ABBA1147 Peptide #
3	18	72.0	59	4	ABBA1147 Peptide #
4	18	72.0	59	4	ABBA1147 Peptide #
5	18	72.0	59	4	ABBA1147 Peptide #
6	18	72.0	59	4	ABBA1147 Peptide #
7	18	72.0	59	4	ABBA1147 Peptide #
8	18	72.0	59	4	ABBA1147 Peptide #
9	18	72.0	59	4	ABBA1147 Peptide #
10	18	72.0	59	4	ABBA1147 Peptide #
11	18	72.0	59	4	ABBA1147 Peptide #
12	18	72.0	59	4	ABBA1147 Peptide #
13	18	72.0	59	4	ABBA1147 Peptide #
14	18	72.0	59	4	ABBA1147 Peptide #
15	18	72.0	59	4	ABBA1147 Peptide #
16	18	72.0	59	4	ABBA1147 Peptide #
17	18	72.0	59	4	ABBA1147 Peptide #
18	18	72.0	59	4	ABBA1147 Peptide #
19	18	72.0	59	4	ABBA1147 Peptide #
20	18	72.0	59	4	ABBA1147 Peptide #
21	18	72.0	59	4	ABBA1147 Peptide #
22	18	72.0	59	4	ABBA1147 Peptide #
23	18	72.0	59	4	ABBA1147 Peptide #
24	18	72.0	59	4	ABBA1147 Peptide #

25	18	72.0	169	7	ABO83337 Pseudomon
26	18	72.0	174	8	ADP30801 Human sec
27	18	72.0	193	7	ABO73977 Pseudomon
28	18	72.0	198	8	ADP30493 Human sec
29	18	72.0	198	8	ADP30492 Human sec
30	18	72.0	198	8	ADP30477 Human sec
31	18	72.0	198	8	ADP30481 Human sec
32	18	72.0	198	8	ADP30491 Human sec
33	18	72.0	228	8	ADP31281 Human sec
34	18	72.0	229	7	ABO80501 Pseudomon
35	18	72.0	233	7	ABO77555 Pseudomon
36	18	72.0	234	8	ADP31468 Human sec
37	18	72.0	242	8	ADP31616 Human sec
38	18	72.0	246	5	AAU75930 Androgen
39	18	72.0	246	5	AAU75930 Androgen
40	18	72.0	249	8	ADP30754 Human sec
41	18	72.0	254	8	ADP31396 Human sec
42	18	72.0	279	8	ADP30846 Human sec
43	18	72.0	279	8	ADP31523 Human sec
44	18	72.0	279	8	ADP30844 Human sec
45	18	72.0	281	4	ADP48050 Signal tr
46	18	72.0	281	8	ADP43729 Corn DBE
47	18	72.0	286	7	ABO79000 Pseudomon
48	18	72.0	297	8	ADP31192 Human sec
49	18	72.0	312	8	ADP31454 Human sec
50	18	72.0	317	8	ADP31615 Human sec
51	18	72.0	339	8	ADP30519 Human sec
52	18	72.0	346	8	ADP31355 Human sec
53	18	72.0	348	8	ADP31441 Human sec
54	18	72.0	348	8	ADP30985 Human sec
55	18	72.0	348	8	ADP31676 Human sec
56	18	72.0	354	8	ADP30783 Human sec
57	18	72.0	354	8	ADP30779 Human sec
58	18	72.0	354	8	ADP30784 Human sec
59	18	72.0	354	8	ADP30778 Human sec
60	18	72.0	360	8	ADP31209 Human sec
61	18	72.0	363	8	ADP30659 Human sec
62	18	72.0	369	8	ADP30582 Human sec
63	18	72.0	372	8	ADP31492 Human sec
64	18	72.0	372	8	ADP31610 Human sec
65	18	72.0	382	8	ADP31221 Human sec
66	18	72.0	387	8	ADP31368 Human sec
67	18	72.0	396	8	ADP31378 Human sec
68	18	72.0	398	7	ABO71161 Pseudomon
69	18	72.0	405	8	ADP31534 Human sec
70	18	72.0	411	8	ADP30729 Human sec
71	18	72.0	426	8	ADP31495 Human sec
72	18	72.0	442	8	ADP31134 Human sec
73	18	72.0	444	8	ADP31640 Human sec
74	18	72.0	471	8	ADP31083 Human sec
75	18	72.0	474	8	ADP30852 Human sec
76	18	72.0	483	8	ADP30853 Human sec
77	18	72.0	483	8	ADP31628 Human sec
78	18	72.0	489	8	ADP31380 Human sec
79	18	72.0	519	8	ADP31190 Human sec
80	18	72.0	525	8	ADP31227 Human sec
81	18	72.0	528	8	ADP30529 Human sec
82	18	72.0	528	8	ADP31525 Human sec
83	18	72.0	533	8	ADP31279 Human sec
84	18	72.0	543	8	ADP30872 Human sec
85	18	72.0	549	8	ADP30855 Human sec
86	18	72.0	552	8	ADP31524 Human sec
87	18	72.0	552	8	ADP31254 Human sec
88	18	72.0	558	8	ADP31255 Human sec
89	18	72.0	558	8	ADP30917 Human sec
90	18	72.0	592	8	ADP31645 Human sec
91	18	72.0	603	8	ADP31264 Human sec
92	18	72.0	610	8	ADP31169 Human sec
93	18	72.0	612	8	ADP31132 Human sec
94	18	72.0	615	8	ADP31361 Human sec
95	18	72.0	615	8	ADP30803 Human sec
96	18	72.0	615	8	ADP31360 Human sec
97	18	72.0	615	8	ADP31360 Human sec

98	18	72.0	637	8	ADP31397	ADp31397 Human sec	171	18	72.0	1092	8	ADP31358	ADp31358 Human sec
99	18	72.0	639	8	ADP30861	Adp30861 Human sec	172	18	72.0	1107	8	ADP30504	ADp30504 Human sec
100	18	72.0	639	9	ADZ10464	AdZ10464 P. gingiv	173	18	72.0	1110	8	ADP31430	ADp31430 Human sec
101	18	72.0	642	8	ADP30748	Adp30748 Human sec	174	18	72.0	1116	8	ADP31692	ADp31692 Human sec
102	18	72.0	660	8	ADP31606	Adp31606 Human sec	175	18	72.0	1128	8	ADP30483	ADp30483 Human sec
103	18	72.0	665	8	ADP30571	Adp30571 Human sec	176	18	72.0	1133	8	ADP30931	ADp30931 Human sec
104	18	72.0	681	8	ADP30868	Adp30868 Human sec	177	18	72.0	1133	8	ADP30928	ADp30928 Human sec
105	18	72.0	681	8	ADP31582	Adp31582 Human sec	178	18	72.0	1133	8	ADP30933	ADp30933 Human sec
106	18	72.0	687	8	ADP31364	Adp31364 Human sec	179	18	72.0	1133	8	ADP30939	ADp30939 Human sec
107	18	72.0	699	8	ADP30747	Adp30747 Human sec	180	18	72.0	1165	8	ADP30879	ADp30879 Human sec
108	18	72.0	705	8	ADP31623	Adp31623 Human sec	181	18	72.0	1167	8	ADP30502	ADp30502 Human sec
109	18	72.0	708	8	ADP31191	Adp31191 Human sec	182	18	72.0	1168	8	ADP31046	ADp31046 Human sec
110	18	72.0	711	8	ADP31440	Adp31440 Human sec	183	18	72.0	1170	8	ADP30919	ADp30919 Human sec
111	18	72.0	729	8	ADP30541	Adp30541 Human sec	184	18	72.0	1170	8	ADP30922	ADp30922 Human sec
112	18	72.0	732	8	ADP30650	Adp30650 Human sec	185	18	72.0	1184	7	ADP06553	ADp06553 Human sec
113	18	72.0	732	8	ADP30650	Adp31010 Human sec	186	18	72.0	1189	8	ADP31043	ADp31043 Human sec
114	18	72.0	738	8	ADP31288	Adp31288 Human sec	187	18	72.0	1191	8	ADP30993	ADp30993 Human sec
115	18	72.0	750	8	ADP30857	Adp30857 Human sec	188	18	72.0	1192	8	ADP30964	ADp30964 Human sec
116	18	72.0	750	8	ADP31131	Adp31131 Human sec	189	18	72.0	1199	8	ADP31044	ADp31044 Human sec
117	18	72.0	750	8	ADP30856	Adp31039 Human sec	190	18	72.0	1200	8	ADP31603	ADp31603 Human sec
118	18	72.0	756	8	ADP31039	Adp31039 Human sec	191	18	72.0	1205	8	ADP30935	ADp30935 Human sec
119	18	72.0	762	8	ADP31012	Adp31012 Human sec	192	18	72.0	1215	8	ADP30898	ADp30898 Human sec
120	18	72.0	768	8	ADP30912	Adp30912 Human sec	193	18	72.0	1215	8	ADP30900	ADp30900 Human sec
121	18	72.0	768	8	ADP31362	Adp31362 Human sec	194	18	72.0	1221	8	ADP31293	ADp31293 Human sec
122	18	72.0	768	8	ADP31363	Adp31363 Human sec	195	18	72.0	1224	8	ADP31426	ADp31426 Human sec
123	18	72.0	771	8	ADP31469	Adp31469 Human sec	196	18	72.0	1227	8	ADP31210	ADp31210 Human sec
124	18	72.0	774	8	ADP30885	Adp31088 Human sec	197	18	72.0	1233	8	ADP30523	ADp30523 Human sec
125	18	72.0	780	8	ADP31077	Adp31077 Human sec	198	18	72.0	1237	2	AAE313791	AAE313791 Fruit fly
126	18	72.0	795	8	ADP31332	Adp31332 Human sec	199	18	72.0	1237	6	AAE30117	AAE30117 Fruit fly
127	18	72.0	816	8	ADP31245	Adp31245 Human sec	200	18	72.0	1260	8	ADP31533	ADp31533 Human sec
128	18	72.0	828	8	ADP31581	Adp31581 Human sec	201	18	72.0	1284	4	ABE871869	ABE871869 D. melano
129	18	72.0	828	8	ADP30878	Adp30878 Human sec	202	18	72.0	1284	5	ABE808376	ABE808376 Human sec
130	18	72.0	834	8	ADP31182	Adp31182 Human sec	203	18	72.0	1302	8	ADP31695	ADp31695 Human sec
131	18	72.0	849	8	ADP31184	Adp31184 Human sec	204	18	72.0	1314	8	ADP31197	ADp31197 Human sec
132	18	72.0	849	8	ADP31184	Adp31184 Human sec	205	18	72.0	1339	8	ADP30732	ADp30732 Human sec
133	18	72.0	861	8	ADP31020	Adp31020 Human sec	206	18	72.0	1341	8	ADP30888	ADp30888 Human sec
134	18	72.0	861	8	ADP31021	Adp31021 Human sec	207	18	72.0	1344	8	ADP31211	ADp31211 Human sec
135	18	72.0	865	8	ADP30566	Adp30566 Human sec	208	18	72.0	1350	8	ADP31138	ADp31138 Human sec
136	18	72.0	885	8	ADP31011	Adp31011 Human sec	209	18	72.0	1358	8	ADP30995	ADp30995 Human sec
137	18	72.0	888	8	ADP31127	Adp31127 Human sec	210	18	72.0	1362	8	ADP31181	ADp31181 Human sec
138	18	72.0	888	8	ADP30971	Adp30971 Human sec	211	18	72.0	1365	8	ADP31035	ADp31035 Human sec
139	18	72.0	889	8	ADP31648	Adp31648 Human sec	212	18	72.0	1371	8	ADP31646	ADp31646 Human sec
140	18	72.0	890	8	ADP31059	Adp31059 Human sec	213	18	72.0	1371	8	ADP30876	ADp30876 Human sec
141	18	72.0	892	8	ADP31578	Adp31578 Human sec	214	18	72.0	1380	8	ADP31566	ADp31566 Human sec
142	18	72.0	906	9	ABE53801	ABE53801 Drosophil	215	18	72.0	1387	8	ADP30946	ADp30946 Human sec
143	18	72.0	921	8	ADP31522	Adp31522 Human sec	216	18	72.0	1417	8	ADP31160	ADp31160 Human sec
144	18	72.0	930	8	ADP31444	Adp31444 Human sec	217	18	72.0	1431	8	ADP31609	ADp31609 Human sec
145	18	72.0	936	8	ADP31597	Adp31597 Human sec	218	18	72.0	1435	6	ABU09754	ABU09754 Human sec
146	18	72.0	939	8	ADP30726	Adp30726 Human sec	219	18	72.0	1435	6	ABU090136	ABU090136 Human sec
147	18	72.0	945	8	ADP31238	Adp31238 Human sec	220	18	72.0	1435	6	ABU96438	ABU96438 Human sec
148	18	72.0	945	8	ADP31237	Adp31237 Human sec	221	18	72.0	1435	6	ABU99047	ABU99047 Human sec
149	18	72.0	948	8	ADP30586	Adp30586 Human sec	222	18	72.0	1435	6	ABU98262	ABU98262 Human sec
150	18	72.0	951	8	ADP30943	Adp30943 Human sec	223	18	72.0	1435	6	ABU91968	ABU91968 Human sec
151	18	72.0	960	8	ADP31471	Adp31471 Human sec	224	18	72.0	1435	6	ABU85272	ABU85272 Human sec
152	18	72.0	960	8	ADP31470	Adp31470 Human sec	225	18	72.0	1435	6	ABO00411	ABO00411 Human sec
153	18	72.0	966	8	ADP30745	Adp30745 Human sec	226	18	72.0	1435	6	ABU88962	ABU88962 Human sec
154	18	72.0	990	8	ADP31553	Adp31553 Human sec	227	18	72.0	1435	6	ABO06458	ABO06458 Human sec
155	18	72.0	1002	8	ADP30866	Adp30866 Human sec	228	18	72.0	1435	6	ABU95518	ABU95518 Human sec
156	18	72.0	1023	8	ADP30994	Adp30994 Human sec	229	18	72.0	1435	6	ABU95208	ABU95208 Human sec
157	18	72.0	1032	8	ADP30874	Adp30874 Human sec	230	18	72.0	1435	6	ABU90756	ABU90756 Human sec
158	18	72.0	1032	8	ADP30863	Adp30863 Human sec	231	18	72.0	1435	6	ABU93918	ABU93918 Human sec
159	18	72.0	1033	8	ADP30984	Adp30984 Human sec	232	18	72.0	1435	6	ABU86192	ABU86192 Human sec
160	18	72.0	1035	8	ADP31552	Adp31552 Human sec	233	18	72.0	1435	6	ABU82047	ABU82047 Human sec
161	18	72.0	1044	8	ADP31517	Adp31517 Human sec	234	18	72.0	1435	6	ABU07908	ABU07908 Human sec
162	18	72.0	1056	8	ADP31082	Adp31082 Human sec	235	18	72.0	1435	6	ABU94228	ABU94228 Human sec
163	18	72.0	1057	8	ADP31592	Adp31592 Human sec	236	18	72.0	1435	6	ABO00101	ABO00101 Human sec
164	18	72.0	1059	8	ADP31042	Adp31042 Human sec	237	18	72.0	1435	6	ABU87112	ABU87112 Human sec
165	18	72.0	1066	8	ADP30563	Adp30563 Human sec	238	18	72.0	1435	6	ABU91353	ABU91353 Human sec
166	18	72.0	1076	6	ABU81145	ABU81145 Human PRO	239	18	72.0	1435	6	ABU90446	ABU90446 Human sec
167	18	72.0	1076	6	ABU68845	ABU68845 Human PRO	240	18	72.0	1435	6	ABU97037	ABU97037 Human sec
168	18	72.0	1082	8	ADP30934	Adp30934 Human sec	241	18	72.0	1435	6	ABO05233	ABO05233 Human sec
169	18	72.0	1086	8	ADP31175	Adp31175 Human sec	242	18	72.0	1464	8	ADP31040	ADp31040 Human sec
170	18	72.0	1090	8	ADQ10187	Adq10187 Human pol	243	18	72.0	1480	8	ADP30557	ADp30557 Human sec



390	17	68.0	45	7	ADC84730	Adc84730 MCF-7 bre	463	17	68.0	157	7	ABO81786	AbO81786 Pseudomon
391	17	68.0	48	2	AAy13202	AAy13202 Human sec	464	17	68.0	158	7	ABO75923	ABO75923 Pseudomon
392	17	68.0	48	9	ABN34232	ABn34232 Opium pop	465	17	68.0	161	4	AAm21861	AAm21861 Peptide #
393	17	68.0	49	8	ABO56643	ABO56643 Human gen	466	17	68.0	161	4	ABR44230	ABR44230 Peptide #
394	17	68.0	51	4	AAU14915	AAu14915 Novel bon	467	17	68.0	161	4	ABM38186	ABM38186 Peptide #
395	17	68.0	51	4	AAU50747	AAu50747 Proxionib	468	17	68.0	161	4	ABR27105	ABR27105 Protein #
396	17	68.0	51	6	ABM47266	ABm47266 Proxionib	469	17	68.0	161	4	AAm77967	AAm77967 Human Don
397	17	68.0	52	6	ABM47265	ABm47265 Proxionib	470	17	68.0	161	4	AAm65265	AAm65265 Human bra
398	17	68.0	52	6	ABM47684	ABm47684 Proxionib	471	17	68.0	161	4	ABG59605	ABG59605 Human liv
399	17	68.0	54	6	ABM44785	ABm44785 Proxionib	472	17	68.0	161	5	ABG46980	ABG46980 Human pep
400	17	68.0	54	6	ABM41304	ABm41304 Proxionib	473	17	68.0	164	8	ADX88412	ADX88412 Plant ful
401	17	68.0	58	5	ABP07833	ABp07833 Human ORF	474	17	68.0	165	8	ADX75562	ADX75562 Plant ful
402	17	68.0	61	4	AAm19250	AAm19250 Peptide #	475	17	68.0	165	8	ADP30587	ADP30587 Human sec
403	17	68.0	61	4	ABR38486	ABr38486 Peptide #	476	17	68.0	165	8	ADP31041	ADP31041 Human sec
404	17	68.0	61	4	AAm31981	AAm31981 Peptide #	477	17	68.0	167	9	AEA20070	AEA20070 Novel hum
405	17	68.0	61	4	ABR32648	ABR32648 Protein #	478	17	68.0	168	6	ABJ39095	ABJ39095 Molecule
406	17	68.0	61	4	AAm71647	AAm71647 Human bon	479	17	68.0	168	8	ADM04991	ADM04991 Human pro
407	17	68.0	61	4	AAm59112	AAm59112 Human bra	480	17	68.0	168	8	ADP31099	ADP31099 Human sec
408	17	68.0	61	4	ABG53331	ABg53331 Human liv	481	17	68.0	169	8	ADP30797	ADP30797 Pseudomon
409	17	68.0	61	5	ABG41461	ABg41461 Human pep	482	17	68.0	169	7	ABO83767	ABO83767 Pseudomon
410	17	68.0	68	4	AAm24008	AAm24008 Human EST	483	17	68.0	173	6	AAU64434	AAU64434 Proxionib
411	17	68.0	68	4	ADP31694	ADp31694 Human sec	484	17	68.0	173	6	ABM60953	ABM60953 Proxionib
412	17	68.0	74	2	AAW94646	AAw94646 TNF-R ext	485	17	68.0	173	6	ABR41796	ABR41796 Human DIT
413	17	68.0	74	4	AAAB69198	AAb69198 Human TNF	486	17	68.0	173	7	ABO83965	ABO83965 Pseudomon
414	17	68.0	74	4	ADH86259	ADh86259 Enterococ	487	17	68.0	173	7	AEA20072	AEA20072 Novel hum
415	17	68.0	74	7	ABO79221	ABO79221 Pseudomon	488	17	68.0	175	4	ABR68003	ABR68003 Drosophil
416	17	68.0	78	8	ADH18879	ADh18879 Human cel	489	17	68.0	175	7	ABO80411	ABO80411 Pseudomon
417	17	68.0	78	6	ABM65345	ABm65345 Proxionib	490	17	68.0	177	6	ABP75865	ABP75865 Human sec
418	17	68.0	79	7	ADP35729	ADp35729 Human hep	491	17	68.0	177	8	ADP30755	ADP30755 Human sec
419	17	68.0	81	7	ABO79675	ABO79675 Pseudomon	492	17	68.0	180	8	ADP30820	ADP30820 Human sec
420	17	68.0	84	3	AAy99920	AAy99920 Peptide e	493	17	68.0	180	8	ADP30825	ADP30825 Human sec
421	17	68.0	84	3	ABG26358	ABg26358 Novel hum	494	17	68.0	180	8	ADP31496	ADP31496 Human sec
422	17	68.0	86	7	ADP35728	ADp35728 Human hep	495	17	68.0	180	8	ADP30828	ADP30828 Human sec
423	17	68.0	92	8	ADP31655	ADp31655 Human sec	496	17	68.0	180	8	ADP30821	ADP30821 Human sec
424	17	68.0	92	4	AAU39635	AAu39635 Proxionib	497	17	68.0	180	8	ADP30826	ADP30826 Human sec
425	17	68.0	92	6	ABM36154	ABm36154 Proxionib	498	17	68.0	180	8	ADP30827	ADP30827 Pseudomon
426	17	68.0	93	6	AAU47955	AAu47955 Proxionib	499	17	68.0	181	7	ABO73220	ABO73220 Pseudomon
427	17	68.0	93	6	ABM44474	ABm44474 Proxionib	500	17	68.0	184	7	ABO77341	ABO77341 Pseudomon
428	17	68.0	93	8	ADP30859	ADp30859 Human sec	501	17	68.0	184	7	ABM88295	ABM88295 Rice abio
429	17	68.0	96	3	ADG27138	ADg27138 Zea maye	502	17	68.0	185	9	AEA20071	AEA20071 Novel hum
430	17	68.0	101	4	ABG12358	ABg12358 Novel hum	503	17	68.0	185	8	ADP31109	ADP31109 Human sec
431	17	68.0	103	7	ABO80232	ABO80232 Pseudomon	504	17	68.0	189	8	ADP30641	ADP30641 Human sec
432	17	68.0	107	4	AAU67241	AAu67241 Proxionib	505	17	68.0	191	4	AAU43753	AAU43753 Human sec
433	17	68.0	107	6	ABM63760	ABm63760 Proxionib	506	17	68.0	191	6	ABM40272	ABM40272 Pseudomon
434	17	68.0	108	3	AAAG18584	AAg18584 Arabidops	507	17	68.0	195	8	ADP30590	ADP30590 Human sec
435	17	68.0	108	3	AAAG15332	AAg15332 Arabidops	508	17	68.0	195	8	ADP30696	ADP30696 Human sec
436	17	68.0	109	8	ABBI1092	ABb11092 Human sec	509	17	68.0	195	8	ADP30837	ADP30837 Human sec
437	17	68.0	112	4	AAU45902	AAu45902 Proxionib	510	17	68.0	203	7	ABO75097	ABO75097 Pseudomon
438	17	68.0	112	4	AAU45902	AAu45902 Proxionib	511	17	68.0	204	7	ABM90226	ABM90226 Rice abio
439	17	68.0	112	5	ABM42431	ABm42431 Proxionib	512	17	68.0	204	8	ADP31421	ADP31421 Human sec
440	17	68.0	114	5	ABP06472	ABp06472 Human ORF	513	17	68.0	204	8	ADP31422	ADP31422 Human sec
441	17	68.0	115	3	AAAG4713	AAg4713 Arabidops	514	17	68.0	204	8	ADP31424	ADP31424 Human sec
442	17	68.0	117	3	AAAG38749	AAg38749 Arabidops	515	17	68.0	204	8	ADP30545	ADP30545 Human sec
443	17	68.0	117	3	AD160202	AD160202 Secretd	516	17	68.0	204	8	ADP31420	ADP31420 Human sec
444	17	68.0	123	8	ADP31326	ADp31326 Human sec	517	17	68.0	207	7	ADM03987	ADM03987 Human pro
445	17	68.0	128	4	AAU50194	AAu50194 Proxionib	518	17	68.0	207	7	ABO71507	ABO71507 Pseudomon
446	17	68.0	128	6	ABM46713	ABm46713 Proxionib	519	17	68.0	208	7	ABO76863	ABO76863 Pseudomon
447	17	68.0	133	8	ADP31069	ADp31069 Human sec	520	17	68.0	211	6	ABR41776	ABR41776 Human DIT
448	17	68.0	134	4	ABG23610	ABg23610 Novel hum	521	17	68.0	211	7	ADP31155	ADP31155 Human dia
449	17	68.0	134	8	ADRO9969	ADrO9969 Human pro	522	17	68.0	213	8	ADP31498	ADP31498 Human sec
450	17	68.0	135	8	ADP30703	ADp30703 Human sec	523	17	68.0	213	8	ADP31499	ADP31499 Human sec
451	17	68.0	136	2	AAy29224	AAy29224 Amino aci	524	17	68.0	214	8	ADY06889	ADY06889 Plant ful
452	17	68.0	138	7	ABO75291	ABO75291 Pseudomon	525	17	68.0	219	8	ADP31171	ADP31171 Human sec
453	17	68.0	138	8	ADP31231	ADp31231 Human sec	526	17	68.0	222	8	ADP31379	ADP31379 Human sec
454	17	68.0	140	4	ABG13377	ABg13377 Novel chr	527	17	68.0	225	8	ADP61861	ADP61861 Transcrip
455	17	68.0	140	4	ABG13377	ABg13377 Novel chr	528	17	68.0	225	7	ABO84376	ABO84376 Pseudomon
456	17	68.0	140	7	ADG33298	ADc33298 Human nov	529	17	68.0	230	7	ABO58717	ABO58717 Human gen
457	17	68.0	140	7	ABO77148	ABO77148 Pseudomon	530	17	68.0	237	8	ADY22944	ADY22944 Plant ful
458	17	68.0	151	7	ABO83144	ABO83144 Pseudomon	531	17	68.0	239	2	ADM17135	ADM17135 Eucalyptu
459	17	68.0	152	7	ABO71621	ABO71621 Pseudomon	532	17	68.0	240	8	AAO68827	AAO68827 Thrombo
460	17	68.0	154	7	ABO76866	ABO76866 Pseudomon	533	17	68.0	240	8	ADP31579	ADP31579 Human sec
461	17	68.0	154	7	ABO67900	ABO67900 Pseudomon	534	17	68.0	240	8	ADT07292	ADT07292 Lettuce p
462	17	68.0	154	7	ABO74421	ABO74421 Pseudomon	535	17	68.0	240	8	ADX77706	ADX77706 Plant ful



536	17	68.0	242	2	AAR06828	Aar06828	Thrombomo	609	17	68.0	361	4	ABB11492	Abb11492	Human neu
537	17	68.0	243	8	ADP30983	Adp30983	Human sec	610	17	68.0	366	8	ADP31106	Adp31106	Human sec
538	17	68.0	243	8	ADP31343	Adp31343	Human sec	611	17	68.0	366	8	ADP31670	Adp31670	Human sec
539	17	68.0	243	9	ADM17137	Adm17137	Eucalypta	612	17	68.0	373	8	ADP30950	Adp30950	Human sec
540	17	68.0	249	2	AAR06826	Aar06826	Thrombomo	613	17	68.0	373	5	ABB47351	Abb47351	Listeria
541	17	68.0	249	2	ABO683160	Ab0683160	Pseudomon	614	17	68.0	381	8	ADP30655	Adp30655	Human sec
542	17	68.0	252	8	ADP31485	Adp31485	Human sec	615	17	68.0	382	8	ADP31216	Adp31216	Human sec
543	17	68.0	253	2	AAR06825	Aar06825	Thrombomo	616	17	68.0	384	8	ADP30656	Adp30656	Human sec
544	17	68.0	253	2	AAR04241	Aar04241	Thrombin-	617	17	68.0	388	6	ABU39484	Abu39484	Protein e
545	17	68.0	253	2	AAR88514	Aar88514	Thrombomo	618	17	68.0	390	8	ADP31584	Adp31584	Human sec
546	17	68.0	253	4	ABB08326	Abb08326	Human thr	619	17	68.0	393	8	ADP31345	Adp31345	Human sec
547	17	68.0	254	3	AAy83938	Aay83938	Human thr	620	17	68.0	395	8	ADP30909	Adp30909	Human sec
548	17	68.0	254	3	AAy83937	Aay83937	Human thr	621	17	68.0	395	7	ABO76549	Ab076549	Pseudomon
549	17	68.0	258	8	ADP30597	Adp30597	Human sec	622	17	68.0	399	8	ADP31313	Adp31313	Human sec
550	17	68.0	262	9	ABE43167	Aeb43167	Zea maye	623	17	68.0	402	7	ABO82656	Ab082656	Pseudomon
551	17	68.0	264	8	ADP30788	Adp30788	Human sec	624	17	68.0	402	8	ADP31414	Adp31414	Human sec
552	17	68.0	264	8	ADP31412	Adp31412	Human sec	625	17	68.0	406	7	ABO69985	Ab069985	Pseudomon
553	17	68.0	264	8	ADP31527	Adp31527	Human sec	626	17	68.0	411	8	ADP31104	Adp31104	Human sec
554	17	68.0	267	8	ADP30822	Adp30822	Human sec	627	17	68.0	411	8	ABM84979	Abm84979	Human dila
555	17	68.0	270	8	ADP30500	Adp30500	Human sec	628	17	68.0	412	4	ABUS3134	Abus3134	Human tra
556	17	68.0	273	8	ADP31236	Adp31236	Human sec	629	17	68.0	414	8	ADP31477	Adp31477	Human sec
557	17	68.0	274	8	ADY22941	Ady22941	Plant ful	630	17	68.0	417	8	ADP31432	Adp31432	Human sec
558	17	68.0	275	2	AAR06832	Aar06832	Thrombomo	631	17	68.0	417	9	ADM17788	Adm17788	Pinus rad
559	17	68.0	275	2	AAR94609	Aar94609	Human rec	632	17	68.0	420	8	ADP31349	Adp31349	Human sec
560	17	68.0	276	7	ABO69382	Ab069382	Pseudomon	633	17	68.0	421	7	ABO75404	Ab075404	Pseudomon
561	17	68.0	277	2	AAR06834	Aar06834	Thrombomo	634	17	68.0	421	8	ADP31159	Adp31159	Human sec
562	17	68.0	278	6	AAU45578	Aau45578	Propionib	635	17	68.0	421	8	ADS23895	Ads23895	Bacterial
563	17	68.0	278	6	ABM42097	Abm42097	Propionib	636	17	68.0	423	8	ADP30819	Adp30819	Human sec
564	17	68.0	278	8	ADP31134	Adp31134	Human sec	637	17	68.0	423	8	ADP31323	Adp31323	Human sec
565	17	68.0	279	8	ADP31107	Adp31107	Human sec	638	17	68.0	429	8	ADP31100	Adp31100	Human sec
566	17	68.0	279	8	ADP31024	Adp31024	Human sec	639	17	68.0	430	7	ABO69654	Ab069654	Pseudomon
567	17	68.0	279	8	ADP31489	Adp31489	Human sec	640	17	68.0	435	8	ADP31102	Adp31102	Human sec
568	17	68.0	288	8	ADP31453	Adp31453	Human sec	641	17	68.0	437	8	ADP31403	Adp31403	Human sec
569	17	68.0	289	7	ADT58146	Adt58146	Plant pol	642	17	68.0	440	6	ADA48432	Ada48432	Rice
570	17	68.0	290	7	ABO70728	Ab070728	Pseudomon	643	17	68.0	448	4	ABG20341	Abg20341	Novel hum
571	17	68.0	291	7	ABO68749	Ab068749	Pseudomon	644	17	68.0	450	8	ADP31085	Adp31085	Human sec
572	17	68.0	293	5	ABB92518	Abb92518	Herbicida	645	17	68.0	453	8	ADP31316	Adp31316	Human sec
573	17	68.0	293	4	ADN74129	Adn74129	Thale cre	646	17	68.0	456	8	ADP31636	Adp31636	Human sec
574	17	68.0	294	4	AAU67478	Aau67478	Propionib	647	17	68.0	461	2	AAR45335	Aar45335	Thrombomo
575	17	68.0	294	6	ABM63997	Abm63997	Propionib	648	17	68.0	461	8	ADP31634	Adp31634	Human sec
576	17	68.0	294	7	ADP31473	Adp31473	Human sec	649	17	68.0	462	2	AAR45347	Aar45347	Thrombomo
577	17	68.0	297	8	ABO83844	Ab083844	Pseudomon	650	17	68.0	462	2	AAR45349	Aar45349	Thrombomo
578	17	68.0	299	6	ABU19784	Abu19784	Protein e	651	17	68.0	462	2	AAR45337	Aar45337	Thrombomo
579	17	68.0	300	8	ADP30775	Adp30775	Human sec	652	17	68.0	462	2	AAR45339	Aar45339	Thrombomo
580	17	68.0	309	8	ADP30862	Adp30862	Human sec	653	17	68.0	462	2	AAR45343	Aar45343	Thrombomo
581	17	68.0	309	8	ADP30873	Adp30873	Human sec	654	17	68.0	462	2	AAR45355	Aar45355	Thrombomo
582	17	68.0	310	9	AEA39559	Aea39559	Nematode	655	17	68.0	462	2	AAR45341	Aar45341	Thrombomo
583	17	68.0	312	7	ADP30476	Adp30476	Human sec	656	17	68.0	462	2	AAR45342	Aar45342	Thrombomo
584	17	68.0	315	7	ADD30795	Add30795	Plant yle	657	17	68.0	462	2	AAR45336	Aar45336	Thrombomo
585	17	68.0	315	8	AD141783	Ad141783	Plant tra	658	17	68.0	462	2	AAR45348	Aar45348	Thrombomo
586	17	68.0	316	8	ADX94660	Adx94660	Plant ful	659	17	68.0	462	2	AAR45350	Aar45350	Thrombomo
587	17	68.0	318	8	ADP31135	Adp31135	Human sec	660	17	68.0	462	2	AAR45345	Aar45345	Thrombomo
588	17	68.0	320	8	ADP31607	Adp31607	Human sec	661	17	68.0	462	2	AAR45354	Aar45354	Thrombomo
589	17	68.0	320	8	ADP31649	Adp31649	Human sec	662	17	68.0	462	2	AAR45338	Aar45338	Thrombomo
590	17	68.0	330	8	ADP31050	Adp31050	Human sec	663	17	68.0	462	2	AAR45353	Aar45353	Thrombomo
591	17	68.0	331	1	AAp82847	Aap82847	Sequence	664	17	68.0	462	2	AAR45346	Aar45346	Thrombomo
592	17	68.0	333	8	ADP31442	Adp31442	Human sec	665	17	68.0	462	2	AAR45344	Aar45344	Thrombomo
593	17	68.0	334	9	ABM96161	Abm96161	M. xanthu	666	17	68.0	463	2	AAR45340	Aar45340	Thrombomo
594	17	68.0	338	8	ADL06030	Adl06030	M. catarr	667	17	68.0	466	5	ABB93873	Abb93873	Herbicida
595	17	68.0	339	8	ADP30892	Adp30892	Human sec	668	17	68.0	471	8	ADP31567	Adp31567	Human sec
596	17	68.0	339	8	ADP30702	Adp30702	Human sec	669	17	68.0	472	8	ADP31222	Adp31222	Human sec
597	17	68.0	340	6	ABM66183	Abm66183	Propionib	670	17	68.0	475	2	AAR22032	Aar22032	Truncated
598	17	68.0	344	6	ABO01324	Ab001324	Human pro	671	17	68.0	476	2	AAR78725	Aar78725	Mature th
599	17	68.0	344	7	ADM26628	Adm26628	Hyperther	672	17	68.0	476	2	AAR86377	Aar86377	Modified
600	17	68.0	344	8	ADN96020	Adn96020	Human NOV	673	17	68.0	478	2	AAR86376	Aar86376	Modified
601	17	68.0	345	8	ADP31016	Adp31016	Human sec	674	17	68.0	478	8	ADP31007	Adp31007	Human sec
602	17	68.0	345	8	ADP31683	Adp31683	Human sec	675	17	68.0	480	2	AAR22013	Aar22013	Truncated
603	17	68.0	349	6	ABM64594	Abm64594	Propionib	676	17	68.0	480	8	ADP31484	Adp31484	Human sec
604	17	68.0	350	8	ADX57315	Adx57315	Plant ful	677	17	68.0	484	4	ABB63251	Abb63251	Drosophil
605	17	68.0	353	8	ADP31358	Adp31358	Human sec	678	17	68.0	484	6	AAE30102	Aae30102	Drosophil
606	17	68.0	356	4	ABG04360	Abg04360	Novel hum	679	17	68.0	484	9	ABE53795	Aeb53795	Drosophil
607	17	68.0	357	8	ADP30505	Adp30505	Human sec	680	17	68.0	486	2	AAE13877	Aae13877	Thrombin-
608	17	68.0	360	8	ADP31439	Adp31439	Human sec	681	17	68.0	490	5	AAE17531	Aae17531	Human thr

682	17	68.0	490	5	AAE23032	AAe23032 Human thr	755	17	68.0	564	8	ADP31194	Adp31194 Human sec
683	17	68.0	494	2	AAE24400	AAe24400 Recombina	756	17	68.0	564	9	ADY70520	Ady70520 Human bet
684	17	68.0	494	2	AAE78727	AAe78727 Thrombomo	757	17	68.0	567	8	ADP31395	Adp31395 Human sec
685	17	68.0	494	2	AAE78726	AAe78726 Thrombomo	758	17	68.0	568	4	ABE61334	ABe61334 Drosophi1
686	17	68.0	494	2	AAE74701	AAe74701 Human bug	759	17	68.0	575	1	ABE82070	ABe82070 Human thr
687	17	68.0	497	2	AAE94607	AAe94607 Human rec	760	17	68.0	575	2	AAE11534	AAe11534 Human thr
688	17	68.0	497	2	AAE95920	AAe95920 IRTM prot	761	17	68.0	575	2	AAE14341	AAe14341 Human thr
689	17	68.0	498	2	AAE84185	AAe84185 Human der	762	17	68.0	575	2	AAE20639	AAe20639 Human urt
690	17	68.0	498	2	AAE01600	AAe01600 Thrombomo	763	17	68.0	575	2	AAE22189	AAe22189 Sequence
691	17	68.0	498	2	AAE67402	AAe67402 Thrombomo	764	17	68.0	575	2	AAE31572	AAe31572 Human thr
692	17	68.0	500	3	AAE69530	AAe69530 Human sug	765	17	68.0	575	2	AAE41806	AAe41806 Thrombomo
693	17	68.0	502	4	ABG03462	ABg03462 Human thr	766	17	68.0	575	2	AAE43031	AAe43031 Human thr
694	17	68.0	510	8	ADP31549	ADp31549 Human hum	767	17	68.0	575	2	AAE73970	AAe73970 Human thr
695	17	68.0	510	8	ADP31049	ADp31049 Human sec	768	17	68.0	575	5	AAE17521	AAe17521 Human full
696	17	68.0	515	2	AAE10617	AAe10617 Soluble t	769	17	68.0	575	5	AAE23026	AAe23026 Human thr
697	17	68.0	516	2	AAE22017	AAe22017 Human thr	770	17	68.0	575	6	ABG72575	ABg72575 Human thr
698	17	68.0	516	2	AAE22016	AAe22016 Truncated	771	17	68.0	575	6	ABU52408	ABu52408 Human GPC
699	17	68.0	516	2	AAE22018	AAe22018 Human thr	772	17	68.0	575	7	ABM78950	ABm78950 Breast ca
700	17	68.0	516	2	AAE09347	AAe09347 Human thr	773	17	68.0	575	7	ADP54760	ADp54760 Human PRO
701	17	68.0	516	2	AAE09348	AAe09348 Human thr	774	17	68.0	575	7	ADP48168	ADp48168 Human PRO
702	17	68.0	516	3	AAE83934	AAe83934 Human thr	775	17	68.0	575	8	ADL24151	ADl24151 Human NOV
703	17	68.0	516	3	AAE83935	AAe83935 Human thr	776	17	68.0	575	8	ADN04006	ADn04006 Antipsori
704	17	68.0	516	3	AAE95929	AAe95929 Human thr	777	17	68.0	575	8	ADP12616	ADp12616 Protein e
705	17	68.0	516	6	ABU08701	ABu08701 Thrombomo	778	17	68.0	575	8	ADP31143	ADp31143 Human sec
706	17	68.0	516	6	ABU08703	ABu08703 Thrombomo	779	17	68.0	575	8	ADP28681	ADp28681 Human thr
707	17	68.0	516	6	ADP31271	ADa31271 Human sol	780	17	68.0	575	8	ADP79506	ADp79506 Human thr
708	17	68.0	516	7	ADP31275	ADa31275 Human sol	781	17	68.0	575	8	ADP32191	ADp32191 Human thr
709	17	68.0	516	8	ADP30871	ADp30871 Human sec	782	17	68.0	575	8	ADU06331	ADu06331 Novel brc
710	17	68.0	516	8	ADP31418	ADp31418 Human sec	783	17	68.0	575	7	ABO799310	ABo799310 Pseudomon
711	17	68.0	519	6	AAE30101	AAe30101 Drosophi1	784	17	68.0	588	8	ADP30877	ADp30877 Human sec
712	17	68.0	522	8	ADP31081	ADp31081 Human sec	785	17	68.0	595	2	AAE35478	AAe35478 Lymphocyt
713	17	68.0	522	8	ADP31018	ADp31018 Human sec	786	17	68.0	595	2	AAE24017	AAe24017 Human sol
714	17	68.0	524	4	AAU07370	AAu07370 G protei	787	17	68.0	595	2	AAE30936	AAe30936 Human CD3
715	17	68.0	531	8	ADP31540	ADp31540 Human sec	788	17	68.0	595	4	AAE50519	AAe50519 Human tum
716	17	68.0	531	8	ADP31696	ADp31696 Human sec	789	17	68.0	595	5	AAU78088	AAu78088 Human CD3
717	17	68.0	532	9	AAE52404	AAe52404 Human thr	790	17	68.0	595	6	ABP97381	ABp97381 Human CD3
718	17	68.0	533	1	AAE80641	AAe80641 Nemacode	791	17	68.0	595	6	ABU04173	ABu04173 Human exp
719	17	68.0	544	8	ADP06563	ADp06563 Plant ful	792	17	68.0	595	6	ABU04172	ABu04172 Human exp
720	17	68.0	544	8	ADP31639	ADp31639 Human sec	793	17	68.0	595	6	ABU04178	ABu04178 Human exp
721	17	68.0	549	8	ADP44055	ADp44055 Human ple	794	17	68.0	595	6	ABU04176	ABu04176 Human exp
722	17	68.0	551	8	ADP30540	ADp30540 Human sec	795	17	68.0	595	6	ABU04177	ABu04177 Human exp
723	17	68.0	552	8	ADP30875	ADp30875 Human sec	796	17	68.0	595	6	ABU04175	ABu04175 Human exp
724	17	68.0	552	8	ADP30875	ADp30875 Human sec	797	17	68.0	595	6	ABU04179	ABu04179 Human exp
725	17	68.0	554	5	AAE17522	AAe17522 Human thr	798	17	68.0	595	6	ABU04170	ABu04170 Human exp
726	17	68.0	554	5	AAE23027	AAe23027 Human thr	799	17	68.0	595	6	ABU04169	ABu04169 Human exp
727	17	68.0	555	8	ADP31168	ADp31168 Human sec	800	17	68.0	595	6	ABP71420	ABp71420 Human Cof
728	17	68.0	555	8	ADP31416	ADp31416 Human sec	801	17	68.0	595	6	ADP19427	ADp19427 Human CD3
729	17	68.0	555	8	ADP31417	ADp31417 Human sec	802	17	68.0	595	6	ADP25553	ADp25553 Binding d
730	17	68.0	557	5	AAE17592	AAe17592 Human thr	803	17	68.0	595	7	ADU24289	ADl24289 Human CD3
731	17	68.0	557	5	AAE17594	AAe17594 Human thr	804	17	68.0	595	8	ADU24289	ADl24289 Human CD3
732	17	68.0	557	5	AAE17524	AAe17524 Human thr	805	17	68.0	595	8	ADQ6799	ADq6799 Human CD3
733	17	68.0	557	5	AAE17528	AAe17528 Human thr	806	17	68.0	595	8	ADQ59520	ADq59520 Human can
734	17	68.0	557	5	AAE17526	AAe17526 Human thr	807	17	68.0	595	8	ADQ59520	ADq59520 Human can
735	17	68.0	557	5	AAE17526	AAe17526 Human thr	808	17	68.0	595	8	ADQ59520	ADq59520 Human can
736	17	68.0	557	5	AAE17523	AAe17523 Human thr	809	17	68.0	595	8	ADP23994	ADp23994 PRO polyp
737	17	68.0	557	5	AAE17523	AAe17523 Human thr	810	17	68.0	595	9	ADY30129	ADy30129 Human CD3
738	17	68.0	557	5	AAE17596	AAe17596 Human thr	811	17	68.0	595	9	ADZ13901	ADz13901 Human can
739	17	68.0	557	5	AAE17591	AAe17591 Human thr	812	17	68.0	600	8	ADP30865	ADp30865 Human sec
740	17	68.0	557	5	AAE17595	AAe17595 Human thr	813	17	68.0	600	8	ADP31150	ADp31150 Human sec
741	17	68.0	557	5	AAE17597	AAe17597 Human thr	814	17	68.0	603	8	ADP31287	ADp31287 Human sec
742	17	68.0	557	5	AAE17525	AAe17525 Human thr	815	17	68.0	603	8	ADP30507	ADp30507 Human sec
743	17	68.0	557	5	AAE17529	AAe17529 Human thr	816	17	68.0	605	8	ADP31229	ADp31229 Human sec
744	17	68.0	557	5	AAE17530	AAe17530 Human thr	817	17	68.0	606	8	ADP31263	ADp31263 Human sec
745	17	68.0	557	5	AAE17598	AAe17598 Human thr	818	17	68.0	609	8	ADP31411	ADp31411 Human sec
746	17	68.0	557	5	AAE17527	AAe17527 Human thr	819	17	68.0	612	8	ADP31064	ADp31064 Human sec
747	17	68.0	557	5	AAE23031	AAe23031 Human thr	820	17	68.0	617	8	ADN24090	ADn24090 Bacterial
748	17	68.0	557	5	AAE23029	AAe23029 Human thr	821	17	68.0	621	8	ADP31147	ADp31147 Human sec
749	17	68.0	557	5	AAE23030	AAe23030 Human thr	822	17	68.0	623	9	AAE39561	AAe39561 Nemacode
750	17	68.0	557	5	AAE23028	AAe23028 Human thr	823	17	68.0	624	8	ADP31324	ADp31324 Human sec
751	17	68.0	557	8	ADP31103	ADp31103 Human sec	824	17	68.0	624	8	ADP31325	ADp31325 Human sec
752	17	68.0	557	8	ADQ65702	ADq65702 Novel hum	825	17	68.0	627	8	ADP31388	ADp31388 Human sec
753	17	68.0	558	8	ADP31257	ADp31257 Human hum	826	17	68.0	629	8	ADP31295	ADp31295 Human sec
754	17	68.0	561	4	ABB70236	ABb70236 Drosophi1	827	17	68.0	631	4	ABG21715	ABg21715 Novel hum

828	17	68.0	631	6	ABG72572	Human	thr	Abg72572	Human	thr	901	17	68.0	813	8	ADP31282	ADp31282	Human	sec	Adp31282	Human	sec
829	17	68.0	639	8	ADP31521	Human	sec	Adp31521	Human	sec	902	17	68.0	813	8	ADP30561	ADp30561	Human	sec	Adp30561	Human	sec
830	17	68.0	642	8	ADP31265	Human	sec	Adp31265	Human	sec	903	17	68.0	814	4	ABB68374	ABb68374	Drosophila		Abb68374	Drosophila	
831	17	68.0	654	8	ADP30656	Human	sec	Adp30656	Human	sec	904	17	68.0	821	8	ADP30679	ADp30679	Human	sec	Adp30679	Human	sec
832	17	68.0	658	8	ADP31226	Human	sec	Adp31226	Human	sec	905	17	68.0	821	8	ADP30680	ADp30680	Human	sec	Adp30680	Human	sec
833	17	68.0	666	8	ADP31547	Human	sec	Adp31547	Human	sec	906	17	68.0	828	8	ADP31569	ADp31569	Human	sec	Adp31569	Human	sec
834	17	68.0	666	8	ADP30867	Human	sec	Adp30867	Human	sec	907	17	68.0	831	8	ADP31303	ADp31303	Human	sec	Adp31303	Human	sec
835	17	68.0	669	8	ADP31142	Human	sec	Adp31142	Human	sec	908	17	68.0	831	8	ADP31333	ADp31333	Human	sec	Adp31333	Human	sec
836	17	68.0	669	8	ADP31598	Human	sec	Adp31598	Human	sec	909	17	68.0	831	8	ADP31123	ADp31123	Human	sec	Adp31123	Human	sec
837	17	68.0	669	8	ADP31493	Human	sec	Adp31493	Human	sec	910	17	68.0	831	8	ADP31179	ADp31179	Human	sec	Adp31179	Human	sec
838	17	68.0	670	8	ADP31001	Human	sec	Adp31001	Human	sec	911	17	68.0	843	8	ADP30663	ADp30663	Human	sec	Adp30663	Human	sec
839	17	68.0	675	8	ADP31438	Human	sec	Adp31438	Human	sec	912	17	68.0	852	8	ADP30969	ADp30969	Human	sec	Adp30969	Human	sec
840	17	68.0	677	8	ADP31585	Human	sec	Adp31585	Human	sec	913	17	68.0	852	8	ADP30664	ADp30664	Human	sec	Adp30664	Human	sec
841	17	68.0	681	8	ADN05602	Antiposori		Adn05602	Antiposori		914	17	68.0	856	8	ADP30972	ADp30972	Human	sec	Adp30972	Human	sec
842	17	68.0	681	8	ADP31053	Human	sec	Adp31053	Human	sec	915	17	68.0	858	4	ABB60597	ABb60597	Drosophila		Abb60597	Drosophila	
843	17	68.0	687	7	ABO77441	Pseudomon		AbO77441	Pseudomon		916	17	68.0	868	8	ADP30736	ADp30736	Human	sec	Adp30736	Human	sec
844	17	68.0	690	8	ADP31686	Human	sec	Adp31686	Human	sec	917	17	68.0	870	8	ADP30646	ADp30646	Human	sec	Adp30646	Human	sec
845	17	68.0	690	8	ADP30750	Human	sec	Adp30750	Human	sec	918	17	68.0	876	8	ADP31220	ADp31220	Human	sec	Adp31220	Human	sec
846	17	68.0	702	8	ADP31518	Human	sec	Adp31518	Human	sec	919	17	68.0	882	8	ADP30487	ADp30487	Human	sec	Adp30487	Human	sec
847	17	68.0	711	8	ADP31215	Human	sec	Adp31215	Human	sec	920	17	68.0	882	8	ADP31688	ADp31688	Human	sec	Adp31688	Human	sec
848	17	68.0	711	8	ADP31652	Human	sec	Adp31652	Human	sec	921	17	68.0	885	8	ADP31198	ADp31198	Human	sec	Adp31198	Human	sec
849	17	68.0	711	8	ADP31535	Human	sec	Adp31535	Human	sec	922	17	68.0	887	8	ADP30554	ADp30554	Human	sec	Adp30554	Human	sec
850	17	68.0	714	8	ADP31561	Human	sec	Adp31561	Human	sec	923	17	68.0	887	8	ADP30548	ADp30548	Human	sec	Adp30548	Human	sec
851	17	68.0	720	5	AAO14994	Laminin-1		AaO14994	Laminin-1		924	17	68.0	891	8	ADP31230	ADp31230	Human	sec	Adp31230	Human	sec
852	17	68.0	720	5	ADW47628	Laminin-5		AdW47628	Laminin-5		925	17	68.0	894	8	ADP30851	ADp30851	Human	sec	Adp30851	Human	sec
853	17	68.0	725	8	ADP31092	Human	sec	Adp31092	Human	sec	926	17	68.0	900	8	ADP31337	ADp31337	Human	sec	Adp31337	Human	sec
854	17	68.0	727	7	ABM88246	Rice	abio	Abm88246	Rice	abio	927	17	68.0	908	6	ABU08492	ABu08492	Alpha-hel		Abu08492	Alpha-hel	
855	17	68.0	728	8	ADP30508	Human	sec	Adp30508	Human	sec	928	17	68.0	918	8	ADP31459	ADp31459	Human	sec	Adp31459	Human	sec
856	17	68.0	731	4	ABG29843	Novel	hum	Abg29843	Novel	hum	929	17	68.0	922	8	ADP30546	ADp30546	Human	sec	Adp30546	Human	sec
857	17	68.0	735	4	ADP31520	Human	sec	Adp31520	Human	sec	930	17	68.0	925	5	AAO14246	AAO14246	Human	pre	AAO14246	Human	pre
858	17	68.0	739	4	ABR70356	Drosophila		Abb70356	Drosophila		931	17	68.0	928	6	ABU20097	ABu20097	Protein	e	Abu20097	Protein	e
859	17	68.0	739	4	ADP31196	Human	sec	Adp31196	Human	sec	932	17	68.0	933	8	ADP31140	ADp31140	Human	sec	Adp31140	Human	sec
860	17	68.0	745	8	ADP30982	Human	sec	Adp30982	Human	sec	933	17	68.0	936	8	ADP31568	ADp31568	Human	sec	Adp31568	Human	sec
861	17	68.0	753	8	ADP30988	Human	sec	Adp30988	Human	sec	934	17	68.0	936	8	ADP31542	ADp31542	Human	sec	Adp31542	Human	sec
862	17	68.0	755	8	ADP31559	Human	sec	Adp31559	Human	sec	935	17	68.0	939	8	ADP31541	ADp31541	Human	sec	Adp31541	Human	sec
863	17	68.0	757	8	ADP30925	Human	sec	Adp30925	Human	sec	936	17	68.0	947	8	ADP30937	ADp30937	Human	sec	Adp30937	Human	sec
864	17	68.0	759	8	ADP31141	Human	sec	Adp31141	Human	sec	937	17	68.0	950	8	ADP31167	ADp31167	Human	sec	Adp31167	Human	sec
865	17	68.0	762	8	ADP30891	Human	sec	Adp30891	Human	sec	938	17	68.0	957	8	ADP31528	ADp31528	Human	sec	Adp31528	Human	sec
866	17	68.0	765	8	ADP31149	Human	sec	Adp31149	Human	sec	939	17	68.0	981	8	ADP30547	ADp30547	Human	sec	Adp30547	Human	sec
867	17	68.0	768	2	AAK27684	Human	bet	AaK27684	Human	bet	940	17	68.0	992	8	ADP31057	ADp31057	Human	sec	Adp31057	Human	sec
868	17	68.0	769	5	ABG66351	Human	ova	Abg66351	Human	ova	941	17	68.0	996	8	ADP31538	ADp31538	Human	sec	Adp31538	Human	sec
869	17	68.0	769	6	ABU56723	Lung	canc	Abu56723	Lung	canc	942	17	68.0	1002	8	ADP39639	ADp39639	Human	pan	Adg39639	Human	pan
870	17	68.0	769	7	ADB80492	Ovarian	c	AdB80492	Ovarian	c	943	17	68.0	1008	8	ADP30721	ADp30721	Human	sec	Adp30721	Human	sec
871	17	68.0	769	7	ADN39527	Cancer/an		Adn39527	Cancer/an		944	17	68.0	1010	8	ADP31296	ADp31296	Human	sec	Adp31296	Human	sec
872	17	68.0	769	7	ADN39162	Cancer/an		Adn39162	Cancer/an		945	17	68.0	1017	8	ADP31268	ADp31268	Human	sec	Adp31268	Human	sec
873	17	68.0	769	7	ADN39597	Cancer/an		Adn39597	Cancer/an		946	17	68.0	1030	8	ADP30913	ADp30913	Human	sec	Adp30913	Human	sec
874	17	68.0	769	7	ADN86616	Human	int	Adn86616	Human	int	947	17	68.0	1038	8	ADP30860	ADp30860	Human	sec	Adp30860	Human	sec
875	17	68.0	771	8	ADP31244	Human	sec	Adp31244	Human	sec	948	17	68.0	1041	8	ADP30998	ADp30998	Human	sec	Adp30998	Human	sec
876	17	68.0	772	8	ADP30936	Human	sec	Adp30936	Human	sec	949	17	68.0	1048	8	ADP31642	ADp31642	Human	sec	Adp31642	Human	sec
877	17	68.0	774	8	ADP30506	Human	sec	Adp30506	Human	sec	950	17	68.0	1050	8	ADP31376	ADp31376	Human	sec	Adp31376	Human	sec
878	17	68.0	774	8	ADP31373	Human	sec	Adp31373	Human	sec	951	17	68.0	1053	8	ADP30886	ADp30886	Human	sec	Adp30886	Human	sec
879	17	68.0	774	8	ADP31225	Human	sec	Adp31225	Human	sec	952	17	68.0	1064	7	ABO68762	ABO68762	Pseudomon		AbO68762	Pseudomon	
880	17	68.0	775	7	ABM85895	Rice	abio	Abm85895	Rice	abio	953	17	68.0	1065	8	ADP30973	ADp30973	Human	sec	Adp30973	Human	sec
881	17	68.0	779	8	ADP30915	Human	sec	Adp30915	Human	sec	954	17	68.0	1065	8	ADP31347	ADp31347	Human	sec	Adp31347	Human	sec
882	17	68.0	779	8	ADP30897	Human	sec	Adp30897	Human	sec	955	17	68.0	1065	8	ADP31482	ADp31482	Human	sec	Adp31482	Human	sec
883	17	68.0	780	2	AAK06842	Thrombomo		AaK06842	Thrombomo		956	17	68.0	1065	8	ADP31287	ADp31287	Human	sec	Adp31287	Human	sec
884	17	68.0	782	8	ADP30770	Human	sec	Adp30770	Human	sec	957	17	68.0	1070	8	ADP30970	ADp30970	Human	sec	Adp30970	Human	sec
885	17	68.0	782	8	ADP30901	Human	sec	Adp30901	Human	sec	958	17	68.0	1086	8	ADP31447	ADp31447	Human	sec	Adp31447	Human	sec
886	17	68.0	783	8	ADP31284	Human	sec	Adp31284	Human	sec	959	17	68.0	1099	8	ADP30574	ADp30574	Human	sec	Adp30574	Human	sec
887	17	68.0	783	8	ADP31436	Human	sec	Adp31436	Human	sec	960	17	68.0	1104	8	ADP31156	ADp31156	Human	sec	Adp31156	Human	sec
888	17	68.0	788	8	ADP31398	Human	sec	Adp31398	Human	sec	961	17	68.0	1119	8	ADP31452	ADp31452	Human	sec	Adp31452	Human	sec
889	17	68.0	791	4	ABBS8225	Drosophila		Abb8225	Drosophila		962	17	68.0	1125	8	ADP30920	ADp30920	Human	sec	Adp30920	Human	sec
890	17	68.0	796	8	ADP32123	PRO	polyp	Adp32123	PRO	polyp	963	17	68.0	1128	6	ADA15725	ADA15725	C. elegans		Ada15725	C. elegans	
891	17	68.0	799	2	AAW02194	Human	int	AaW02194	Human	int	964	17	68.0	1134	8	ADP30647	ADp30647	Human	sec	Adp30647	Human	sec
892	17	68.0	799	5	AAU76337	Human	ant	AaU76337	Human	ant	965	17	68.0	1134	8	ADP31537	ADp31537	Human	sec	Adp31537		

```
974 17 68.0 1173 8 ADP31155 Human sec
975 17 68.0 1179 8 ADP30486 Human sec
976 17 68.0 1179 8 ADP30577 Human sec
977 17 68.0 1183 8 ADP30536 Human sec
978 17 68.0 1192 8 ADP31180 Human sec
979 17 68.0 1194 8 ADP30682 Human sec
980 17 68.0 1197 8 ADP31034 Human sec
981 17 68.0 1197 8 ADP31342 Human sec
982 17 68.0 1200 9 AD276047 Mouse pre
983 17 68.0 1222 8 ADP30501 Human sec
984 17 68.0 1227 8 ADP31602 Human sec
985 17 68.0 1239 8 ADP31297 Human sec
986 17 68.0 1248 8 ADP31346 Human sec
987 17 68.0 1252 8 ADP30678 Human sec
988 17 68.0 1269 8 ADP31382 Human sec
989 17 68.0 1269 8 ADP31580 Human sec
990 17 68.0 1269 8 ADP31381 Human sec
991 17 68.0 1269 8 ADP303456 Human sec
992 17 68.0 1282 8 ADP31328 Human sec
993 17 68.0 1289 8 ADP30675 Human sec
994 17 68.0 1300 6 ABU88254 Human sec
995 17 68.0 1300 6 ABU90133 Human sec
996 17 68.0 1300 6 ABU96435 Human sec
997 17 68.0 1300 6 ABU99044 Human sec
998 17 68.0 1300 6 ABU98259 Human sec
999 17 68.0 1300 6 ABU91965 Human sec
1000 17 68.0 1300 6 ABU85269 Human sec
```

## ALIGNMENTS

```
RESULT 1
ID ABB41147 standard; peptide; 59 AA.
XX
AC ABB41147;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #8653 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human foetal liver.
XX
PS Claim 27; SEQ ID NO 33782; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
```

CC single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequence is a peptide encoded by a single exon nucleic acid probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pcc\\_sequences](http://wipo.int/pub/published_pcc_sequences)

XX Sequence 59 AA;

SQ

Query Match 72.0%; Score 18; DB 4; Length 59;

Best Local Similarity 22.2%; Pred. No. 2.8e+02; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7;

QY 1 CXXXXXXX 9

DB 13 CXXXXXXX 21

RESULT 2

ID AAM34923 standard; protein; 59 AA.

XX AAM34923;

DT 17-OCT-2001 (first entry)

DE Peptide #8960 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000663.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488897/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human placenta.

XX PS Claim 27; SEQ ID NO 35192; 654pp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENP;

XX CC see AAI13135-AI157546). The present sequence is a peptide encoded by one

XX CC such probe. The probes are useful for producing a microarray for

XX CC predicting, measuring and displaying gene expression in samples derived

XX CC from human placenta. The probes are useful for antenatal diagnosis of

XX CC human genetic disorders

SQ Sequence 59 AA;

Query Match 72.0%; Score 18; DB 4; Length 59;

Best Local Similarity 22.2%; Pred. No. 2.8e+02; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 7;

QY 1 CXXXXXXX 9

Db 13 CSSSSTSTC 21

## RESULT 3

ID ABB25185 standard; protein; 59 AA.

AC ABB25185;

DT 23-JAN-2002 (first entry)

DE Protein #7184 encoded by probe for measuring heart cell gene expression.

KW Human; gene expression; heart; microarray; vascular system;

KM cardiovascular disease; hypertension; cardiac arrhythmia;

KW congenital heart disease.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000666.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488990/53.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX hearts.

XX Claim 15; SEQ ID NO 26955; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for

XX measuring human gene expression in a sample derived from human heart (see

XX ABA2153-ABA41305). The present sequence is a protein encoded by one such

XX probe. The probes may be used for predicting, measuring and displaying

XX gene expression in samples derived from the human heart via microarrays.

XX By measuring gene expression, the probes are useful for predicting,

XX diagnosing, grading, staging, monitoring and prognosing diseases of the

XX human heart and vascular system e.g. cardiovascular disease,

XX hypertension, cardiac arrhythmias and congenital heart disease. Note: The

XX sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 59 AA;

Query Match 72.0%; Score 18; DB 4; Length 59;

Best Local Similarity 22.2%; Pred. No. 2.8e+02;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 13 CSSSSTSTC 21

## RESULT 4

ID AAM74807 standard; protein; 59 AA.

AC AAM74807;

XX 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 35113.

KW Human; bone marrow expressed exon; gene expression analysis; probe;

KM microarray; cancer; leukemia; lymphoma; myeloma.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000666.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488990/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human bone marrow.

XX Example 4; SEQ ID NO 35113; 658pp + Sequence listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukemia and myeloma. The present sequence is a

XX protein encoded by one of the probes of the invention

SQ Sequence 59 AA;

Query Match 72.0%; Score 18; DB 4; Length 59;

Best Local Similarity 22.2%; Pred. No. 2.8e+02;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 13 CSSSSTSTC 21

## RESULT 5

ID AAM62003 standard; protein; 59 AA.

AC AAM62003;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 34108.

KW Human; brain expressed exon; gene expression analysis; probe; microarray;

KM Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000667.

XX 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains.  
XX  
PS Example 4; SEQ ID NO 34108; 650bp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention  
XX  
SQ Sequence 59 AA;

Query Match 72.0%; Score 18; DB 4; Length 59;

Best Local Similarity 22.2%; Pred. No. 2.8e+02; 7; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 9  
Db 13 CXXXXXXXC 21

RESULT 6  
ID ABG56589 standard; peptide; 59 AA.  
XX  
AC ABG56589;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human liver peptide, SEQ ID No 35237.  
XX  
KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
KW hypercholesterolaemia; coronary heart disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200157273-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000664.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human adult liver.  
XX  
PS Claim 27; SEQ ID NO 35237; 658bp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (1) for  
CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 1109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult liver.  
CC (1) may be used for predicting, measuring and displaying gene expression  
CC in samples derived from human adult liver. The genes identified may be  
CC involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
CC associated with coronary heart disease. ABG47348-ABG59930 represent human  
CC liver single exon encoded peptides of the invention. Note: The sequence  
CC information for this patent does not appear in the printed specification  
CC but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 59 AA;

Query Match 72.0%; Score 18; DB 4; Length 59;

Best Local Similarity 22.2%; Pred. No. 2.8e+02; 7; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 9  
Db 13 CXXXXXXXC 21

RESULT 7  
ID ABG44598 standard; peptide; 59 AA.  
XX  
AC ABG44598;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human peptide encoded by genome-derived single exon probe SEQ ID 34263.  
XX  
KW Human; single exon probe; asthma; lung cancer; COPD; IHD;  
KW chronic obstructive pulmonary disease; interstitial lung disease;  
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;  
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KW primary ciliary dyskinesia; pulmonary hypertension;  
KW hyaline membrane disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200186003-A2.  
XX  
PD 15-NOV-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000665.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples.  
XX

PS Claim 27; SEQ ID NO 34263; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of probes  
CC; the novel set of probes which hybridize at high stringency to a nucleic  
CC acid expressed in the human lung; measuring gene expression in a sample  
CC delivered from human lung, comprising (a) contacting the array with a  
CC collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of the  
CC array; identifying exons in a eukaryotic genome, comprising (a)  
CC algorithmically predicting at least one exon from genomic sequences of  
CC the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarray having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene expression  
CC analysis, and for identifying exons in a gene, particularly using human  
CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
CC Kargener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
CC present sequence is a peptide/protein encoded by a single exon probe of  
CC the invention. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pcr\_sequences  
XX

SQ Sequence 59 AA;

Query Match 72.0%; Score 18; DB 5; Length 59;  
Best Local Similarity 22.2%; Pred. No. 2.8e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
DB 13 CXXXXSTSTC 21

RESULT 8  
AAB21190  
ID AAB21190 standard; protein; 108 AA.

AC AAB21190;

DT 08-JUN-2001 (first entry)

XX Exo14 partial protein.

XX Exocytosis pathway protein; Exo protein; antiparkinsonian; antidiabetic;  
XX anti-allergic; antiallergic; antiparkinsonian; anticonvulsant;  
XX vulnerable; asthma; inflammation; allergy; Chediak-Higashi syndrome; CHS;  
XX Alzheimer's disease; Parkinson's disease; Huntington's disease; diabetes;  
XX digestion disorder; wound healing disorder; gene therapy.

OS Mus sp.

XX WO200043419-A2.

XX 27-UTL-2000.

PF 20-JAN-2000; 2000WO-US001431.

PR 20-JAN-1999; 99US-0116534P.

PR 26-JAN-1999; 99US-0117274P.

PR 26-JAN-1999; 99US-0117308P.

PR 26-JAN-1999; 99US-0117312P.

PR 01-FEB-1999; 99US-0118177P.

PR 01-FEB-1999; 99US-0118178P.

PR 09-FEB-1999; 99US-0118286P.

PR 11-FEB-1999; 99US-0119759P.

PR 11-FEB-1999; 99US-0119998P.

XX (RIGEL-) RIGEL PHARM INC.

XX Luo Y;

XX WPI; 2000-482908/42.

DR N-PSDB; AAA89569.

XX New nucleic acids encoding Exo proteins which are useful in the

PT diagnosis, treatment or prevention of exocytosis-mediated disorders such

PT as asthma, inflammation and allergies.

XX Disclosure; Page 149-150; 305pp; English.

XX The present sequence is a polypeptide which is associated with the  
XX exocytosis pathway. CDNA molecules encoding proteins involved in  
XX exocytosis have been isolated by yeast one-hybrid and two-hybrid  
XX screening. Novel proteins, termed Exo proteins, have been identified that  
XX interact with known exocytosis-associated proteins such as GS27, alpha  
XX snap, unc18-1, vamps, snap-23, and the rab family of proteins. Exo  
XX proteins and their agonists and antagonists are useful in the diagnosis,  
XX treatment or prevention of exocytosis-mediated disorders such as asthma,  
XX inflammation, allergies, Chediak-Higashi Syndrome (CHS), Alzheimer's  
XX disease, Parkinson's disease, Huntington's disease, diabetes, digestion  
XX disorders and wound healing disorders. The nucleic acids, antagonists or  
XX agonists of Exo proteins are useful in gene therapy. The nucleic acids  
XX are also useful for generating transgenic or knock-out animals which can  
XX be used in the development and screening of therapeutically useful  
XX reagents

SQ Sequence 108 AA;

Query Match 72.0%; Score 18; DB 3; Length 108;  
Best Local Similarity 22.2%; Pred. No. 3.2e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
DB 33 CXXXXSSSSC 41

RESULT 9  
ABO75321  
ID ABO75321 standard; protein; 134 AA.

AC ABO75321;

DT 29-UTL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #7496.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX

PN	US6551795-B1.
XX	
PD	22-APR-2003.
XX	
PF	18-FEB-1999;
XX	99US-00252991.
PR	18-FEB-1998;
XX	98US-0074788P.
PR	27-JUL-1998;
XX	98US-0094190P.
PA	(GENO-) GENOME THERAPEUTICS CORP.
PI	Rubenfield MJ, Nolling J, DeLoughery C, Bush D;
DR	WPI; 2003-615309/58.
N-P	N-PDSB; ABD08892.
PT	Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, of useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.
PS	Disclosure; SEQ ID NO 24067; 455pp; English.
XX	
CC	The invention relates to Pseudomonas aeruginosa polypeptides and the
CC	polynucleotides encoding them. The sequences are useful in diagnosis and
CC	therapy of pathological conditions, as molecular targets for diagnostics,
CC	prophylaxis and treatment of pathological conditions resulting from a
CC	bacterial infection, for evaluating a compound, such as a polypeptide,
CC	for the ability to bind a P. aeruginosa nucleic acid, as components of
CC	effective antibacterial targets, as targets for antibacterial drugs,
CC	including anti-P. aeruginosa drugs, as templates for recombinant
CC	production of P. aeruginosa-derived peptides or polypeptides, as target
CC	components for diagnosis and/or treatment of P. aeruginosa-caused
CC	infection, and in detection of P. aeruginosa sequences or other sequences
CC	of Pseudomonas species using biochip technology. Sequences ABO67826-
CC	ABO64136 represent P. aeruginosa polypeptides of the invention. Note: The
CC	sequence data for this patent did not form part of the printed
CC	specification but was obtained in electronic format from USPRO at
CC	seqdata.uspto.gov/sequence.html
XX	
SQ	Sequence 134 AA:
QY	Query Match: 72.0%; Score 18; DB 7; Length 134; Best Local Similarity 22.2%; Pred. No. 3.4e+02; Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Db	1 1 CXXXXXXXXX 9 19 CSASTTSC 27
RESULT 10	
ADE72390	ID ADE72390 standard; protein; 139 AA.
XX	ADE72390;
XX	29-JAN-2004 (first entry)
DT	
DE	Human endometrial specific protein, SEQ ID NO 430.
XX	
KV	cytostatic; vaccine; human; endometrial specific genes;
KW	endometrial specific protein; endometrial cancer.
OS	Homo sapiens.
XX	
PN	WO2003060081-A2.
XX	
PD	24-JUL-2003.
XX	
PF	23-DEC-2002; 2002WO-US041612.
XX	
PR	21-DEC-2001; 2001US-0342756P.
XX	

```

PA      (DIAD-) DIADEXUS INC.
XX
XX      Sun Y, Liu C;
XX
XX      WPI; 2003-577666/54.
XX
XX      Nucleic acid molecules and polypeptides useful for diagnosing and
XX      treating endometrial cancer and non-cancerous disease states in
XX      endometrial.
XX
XX      Claim 12; SEQ ID NO 430; 824bp; English.
XX
XX      The invention comprises the amino acid and DNA sequences of human
XX      endometrial specific genes and proteins. The DNA and protein sequences of
XX      the invention are useful for diagnosing, imaging and treating a patient
XX      with endometrial cancer. The present amino acid sequence represents a
XX      human endometrial specific protein of the invention.
XX
XX      Sequence 139 AA;
XX
XX      Query Match          72.0%; Score 18; DB 7; Length 139;
XX      Best Local Similarity 22.2%; Pred. No. 3.4e+02;
XX      Matches      2; Conservative      0; Mismatches      7; Indels      0; Gaps      0
XX
XX      QY      1 CXXXXXXXC 9
XX      DB      60 CSTSSAASC 68
XX
XX      RESULT 11
XX      ADE72391
XX      ID      ADE72391 standard; protein; 139 AA.
XX
XX      ADE72391;
XX      AC
XX      DT      29-JAN-2004 (first entry)
XX
XX      Human endometrial specific protein, SEQ ID NO 431.
XX
XX      cytosolic; vaccine; human; endometrial specific genes;
XX      endometrial specific protein; endometrial cancer.
XX
XX      Homo sapiens.
XX
XX      WO2003060081-A2.
XX
XX      24-JUL-2003.
XX
XX      23-DEC-2002; 2002WO-US041612.
XX
XX      21-DEC-2001; 2001US-0342756P.
XX
XX      (DIAD-) DIADEXUS INC.
XX
XX      Sun Y, Liu C;
XX
XX      WPI; 2003-577666/54.
XX
XX      Nucleic acid molecules and polypeptides useful for diagnosing and
XX      treating endometrial cancer and non-cancerous disease states in
XX      endometrial.
XX
XX      Claim 12; SEQ ID NO 431; 824bp; English.
XX
XX      The invention comprises the amino acid and DNA sequences of human
XX      endometrial specific genes and proteins. The DNA and protein sequences of
XX      the invention are useful for diagnosing, imaging and treating a patient
XX      with endometrial cancer. The present amino acid sequence represents a
XX      human endometrial specific protein of the invention.
XX
XX      Sequence 139 AA;
XX
XX      Query Match          72.0%; Score 18; DB 7; Length 139;
XX      Best Local Similarity 22.2%; Pred. No. 3.4e+02;
XX      Matches      2; Conservative      0; Mismatches      7; Indels      0; Gaps      0

```



Best Local Similarity 22.24; Pred. No. 3.4e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
Db 60 CSTSSAASC 68

## RESULT 12

AAM93094  
ID AAM93094 standard; protein; 140 AA.

XX AAM93094;

DT 06-NOV-2001 (first entry)

XX Human digestive system antigen SEQ ID NO: 2443.

XX Human; digestive system antigen; gene therapy; cancer; appendicitis;

KM ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;

KM digestive system disorder; Meckel's diverticulum.

OS Homo sapiens.

PN WO200155314-A2.

PD 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001324.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0217496P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-022679P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229333P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 12-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234423P.

PR 21-SEP-2000; 2000US-0234474P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235634P.

PR 27-SEP-2000; 2000US-0235636P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241212P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0244674P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246539P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX MPI: 2001-502630/55.  
DR N-PSDB; AAK88867.  
XX  
XX  
PT Polynucleotides encoding digestive system antigens, useful for  
PT diagnosing, treating, preventing and/or prognosing disorders of the  
PT digestive system, particularly cancer and cancer metastases.  
XX  
XX  
PS Claim 11; SEQ ID NO 2443; 986bp; English.  
XX  
XX  
CC The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic  
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a digestive system antigen of  
CC the invention  
CC  
CC  
SQ Sequence 140 AA;  
  
Query Match 72.0%; Score 18; DB 4; Length 140;  
Best Local Similarity 22.2%; Pred. No. 3.4e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
OY 1 CXXXXXXXC 9  
Db 104 CTAATATAC 112  
  
RESULT 13  
AAU20102  
ID AAU20102 standard; protein; 140 AA.  
XX  
AC AAU20102;  
XX  
DT 06-DEC-2001 (first entry)  
XX  
DE Human liver associated polypeptide #133.  
XX  
XX Liver associated protein; human; mouse; rabbit; goat; horse; cat; dog;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic;  
KW antineumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; nocitropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnerrary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;

KW anti-infertility.  
XX  
OS Homo sapiens.  
XX  
PN WO200155355-A1.  
XX  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001351.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-020515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0229824P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229511P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234233P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-023498P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 29-SEP-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241211P.  
PR 20-OCT-2000; 2000US-0241212P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251899P.  
PR 08-DEC-2000; 2000US-0251900P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI, 2001-457728/49.  
DR N-PSDB; NAB31813.  
XX  
PT Isolated nucleic acid molecule encoding a human liver related protein is  
PT used in preventing, treating or ameliorating disorders of the liver  
PT particularly cancer of the liver.  
XX  
PS Claim 11; SEQ ID NO 289; 526pp; English.  
XX  
XX Sequences AAU19970-AAU20115 represent the liver associated polypeptides  
CC of the invention. Liver associated polypeptides and their associated  
CC polynucleotides are useful in the diagnosis, treatment and prevention of  
CC various types of disorders in e.g. humans, mice, rabbits, goats, horses,  
CC cats, dogs, chickens or sheep. A pathological condition can be determined  
CC by detecting the presence or absence of a mutation in a liver associated  
CC polynucleotide. The treatable disorders include autoimmune diseases such  
CC as rheumatoid arthritis, hyperproliferative disorders such as neoplasms  
CC of the breast or liver, cardiovascular disorders such as cardiac arrest,  
CC cerebrovascular disorders such as cerebral ischaemia, nervous system  
CC disorders such as Alzheimer's disease, infections caused by bacteria,  
CC viruses and fungi, ocular disorders such as corneal infection, endocrine  
CC disorders such as Crohn's disease, renal disorders such as  
CC glomerulonephritis and respiratory disorders such as asthma and pleurisy.  
CC The polypeptides can also be used to aid wound healing, to prevent skin  
CC aging due to sunburn, to maintain organs before transplantation, to  
CC regenerate tissues and in chemotaxis. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

Query Match 72.0%; Score 18; DB 4; Length 140;  
Best Local Similarity 22.2%; Pred. No. 3.4e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
Db 104 CTATSATAC 112

RESULT 14  
ABP40963  
ID ABP40963 standard; protein; 140 AA.  
XX  
AC ABP40963;  
XX  
DT 24-JUL-2002 (first entry)  
XX  
DE Human liver antigen HFLVG70, SEQ ID NO:289.  
XX  
KW Human; liver antigen; liver disorder; hepatic disorder; infection;  
KW hepatitis; viral; parasitic; bacterial; fungal; inflammatory condition;  
KW cirrhosis; granulomatous hepatitis; toxin damage; drug damage;  
KW autoimmune disease; Wilson's disease; primary biliary cirrhosis;  
KW neoplastic disorder; cancer; tumour; portal hypertension;  
KW gastrointestinal disorder; hepatitis; drug screening; gene therapy;  
KW chromosome mapping; forensic analysis; antibody preparation;  
KW hepatocytic; cytostatic; antiinflammatory; vitucide; antibacterial;  
KW fungicide; parasiticide; antidote; immunosuppressive.  
XX  
XX Homo sapiens.  
OS  
XX  
PN US2002042096-A1.  
XX  
XX 11-APR-2002.  
PD  
XX 17-JAN-2001; 2001US-00764887.

XX 31-JAN-2000; 2000US-0179655P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 22-AUG-2000; 2000US-0226688P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 XX (ROSE/) ROSEN C A.  
 PA (RUBB/) RUBEN S M.  
 PA (BARA/) BARASH S C.  
 XX  
 PT Rosen CA, Ruben SM, Barash SC;  
 DR WPI; 2002-381944/41.  
 DR N-PSDB; AEN90168.  
 XX  
 PT New nucleic acid encoding human liver antigens, useful for diagnosis,  
 PT treatment and prevention of e.g. hepatitis and hepatic cancer, also  
 PT related polypeptides and antibodies.  
 XX  
 PS Claim 11; SEQ ID NO 289; 181bp; English.  
 XX  
 CC The invention relates to 145 novel human liver antigens (ABP40831-  
 CC ABP40975) and to cDNAs encoding them (ABN90036-ABN90180), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human liver antigen  
 CC polynucleotides, antibodies against human liver antigens, and the use of

CC liver antigen polynucleotides and polypeptides in diagnosing, treating,  
 CC prognosing or preventing various disorders of the liver. Such conditions  
 CC include viral infections (e.g., cytomegalovirus, Epstein-Barr virus,  
 CC hepatitis A virus, hepatitis B virus and hepatitis C virus), parasitic  
 CC infections (e.g., *Clonorchis sinensis*, *Echinococcus granulosus* and  
 CC *Entamoeba histolytica*), and also bacterial and fungal infections. Other  
 CC disorders that may be treated include inflammatory conditions (e.g.,  
 CC cirrhosis and granulomatous hepatitis), damage caused by drugs or toxins,  
 CC autoimmune diseases (e.g. Wilson's disease, primary biliary cirrhosis),  
 CC neoplastic disorders (e.g., adenomas, haemangiomas and hepatocellular  
 CC carcinoma), portal hypertension, or gastrointestinal disorders (e.g.,  
 CC peptic ulcers, gastritis and peritoneal diseases). Liver antigen  
 CC polypeptides and polynucleotides may also be used in screening for  
 CC compounds which modulate liver antigen expression or activity. The  
 CC polynucleotides may further be used for gene therapy, chromosome mapping,  
 CC in the identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as molecular weight markers or to prepare  
 CC antibodies useful in disease diagnosis, drug targeting and phenotyping.  
 CC The present sequence represents a human liver antigen of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC USPTO at [seqdata.uspro.gov/sequence/](http://seqdata.uspro.gov/sequence/)  
 CC  
 XX  
 SQ Sequence 140 AA;  
 XX  
 Query Match 72.0%; Score 18; DB 5; Length 140;  
 Best Local Similarity 22.2%; Pred. No. 3 4e+02;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXC 9  
 Db 104 CTATSATAC 112  
 XX  
 RESULT 15  
 ADJ15081 standard; protein; 140 AA.  
 XX  
 ADJ15081;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Human liver-related protein - SEQ ID 289.  
 XX  
 KW liver; virocid; fungicide; antibacterial; antiparasitic; hepatotropic;  
 KW antiinflammatory; cyrostatic; litholytic; antihemetic; antiahrtric;  
 KW neuroprotective; antidiabetic; antioesulant; thrombolytic;  
 KW antiatherosclerotic; cardiac; haemostatic; antiarrhythmic;  
 KW ophthalmological; antiatherosclerotic; vasotropic; osteopathic;  
 KW nootropic; antiparkinsonian; anticonvulsant; neuroleptic; vasotropic;  
 KW cytostatic; gynaecological; viral; fungal; bacterial;  
 KW parasitic infection; cirrhosis; Wilson's disease;  
 KW gastrointestinal disorder; pancreatic; gallbladder; immune; blood;  
 KW hyperproliferative; cardiovascular; respiratory; musculoskeletal system;  
 KW neurological; endocrine; reproductive system; developmental; inherited;  
 KW human.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2003077602-A1.  
 XX  
 PD 24-APR-2003.  
 XX  
 PF 14-FEB-2002; 2002US-00073961.  
 XX  
 XX 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.



CC litholytic, antirheumatic, antiarthritic, neuroprotective, antidiabetic,  
CC anticoagulant, thrombolytic, antiarteriosclerotic, cardiac, haemostatic,  
CC antiarrhythmic, ophthalmologic, antiarteriosclerotic, vasotropic,  
CC osteopathic, nootropic, antiparkinsonian, anticonvulsant, neuroleptic,  
CC vasotropic, cytostatic and gynaecological activities. The polypeptides  
CC and polynucleotides of the invention may be useful for diagnosis,  
CC detection, treatment and/or prevention of disorders of the liver such as  
CC viral, fungal, bacterial or parasitic infections, cirrhosis, Wilson's  
CC disease, gastrointestinal disorders, pancreatic disorders, gallbladder  
CC diseases, immune disorders, blood related disorders, hyperproliferative  
CC disorders, cardiovascular disorders, respiratory disorders,  
CC musculoskeletal system disorders, neurological diseases, endocrine  
CC disorders, reproductive system disorders or developmental and inherited  
CC disorders. The current sequence is that of the human liver-related  
CC protein of the invention. The current sequence is not shown within the  
CC specification per se but was obtained electronically from the USPTO web-  
CC site.

Query Match 72.0%; Score 18; DB 7; Length 140;  
Best Local Similarity 22.2%; Pred. No. 3.4e+02; Mismatches 0; Gaps 0;  
Matches 2; Conservative 0; Indels 0;

Oy 1 CXXXXXXC 9  
Db 104 CTATSATAC 112

RESULT 16  
ADP31453  
ID ADP31453 standard; protein; 144 AA.

AC ADP31453;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2220.

XX Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW Cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX WO2004035732-A2.

XX PD 29-APR-2004.

XX 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 29-AUG-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486466P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3451; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
XX antinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present  
XX sequence represents a human secreted protein. The present sequence is  
XX available on WIPOMB and is not in the specification.

XX Sequence 144 AA;

Query Match 72.0%; Score 18; DB 8; Length 144;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02; Mismatches 7; Indels 0; Gaps 0;  
Matches 2; Conservative 0;

OY 1 CXXXXXXXXC 9  
DB 7 CAATTATTC 15

RESULT 17  
ABO83270  
ID ABO83270 standard; protein; 145 AA.

XX ABO83270;

XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #15445.

KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

XX US651795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD16841.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 32016; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the CC polynucleotides encoding them. The sequences are useful in diagnosis and CC therapy of pathological conditions, as molecular targets for diagnostics, CC prophylaxis and treatment of pathological conditions resulting from a CC bacterial infection, for evaluating a compound, such as a polypeptide, CC for the ability to bind a P. aeruginosa nucleic acid, as components of CC effective antibacterial targets, as targets for antibacterial drugs, CC including anti-P. aeruginosa drugs, as templates for recombinant CC production of P. aeruginosa-derived peptides or polypeptides, as target CC components for diagnosis and/or treatment of P. aeruginosa-caused CC infection, and in detection of P. aeruginosa sequences or other sequences CC of Pseudomonas species using biochip technology. Sequences ABO67826- CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The CC sequence data for this patent did not form part of the printed CC specification but was obtained in electronic format from USPTO at CC seqdata.uspto.gov/sequence.html

XX SQ Sequence 145 AA;

Query Match 72.0%; Score 18; DB 7; Length 145;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 CXXXXXXXXC 9  
DB 20 CSATATTC 28

RESULT 18  
ABO79762  
ID ABO79762 standard; protein; 147 AA.

AC ABO79762;  
XX 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #11937.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

OS Pseudomonas aeruginosa.

XX US651795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD13333.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, useful as molecular targets for diagnostics, prophylaxis and treatment of pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 28508; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the CC polynucleotides encoding them. The sequences are useful in diagnosis and CC therapy of pathological conditions, as molecular targets for diagnostics, CC prophylaxis and treatment of pathological conditions resulting from a CC bacterial infection, for evaluating a compound, such as a polypeptide, CC for the ability to bind a P. aeruginosa nucleic acid, as components of CC effective antibacterial targets, as targets for antibacterial drugs, CC including anti-P. aeruginosa drugs, as templates for recombinant CC production of P. aeruginosa-derived peptides or polypeptides, as target CC components for diagnosis and/or treatment of P. aeruginosa-caused CC infection, and in detection of P. aeruginosa sequences or other sequences CC of Pseudomonas species using biochip technology. Sequences ABO67826- CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The CC sequence data for this patent did not form part of the printed CC specification but was obtained in electronic format from USPTO at CC seqdata.uspto.gov/sequence.html

XX SQ Sequence 147 AA;

Query Match 72.0%; Score 18; DB 7; Length 147;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 CXXXXXXXXC 9  
DB 27 CSTSTTAC 35

RESULT 19  
ADP31511  
ID ADP31511 standard; protein; 153 AA.

XX ADP31511;  
XX 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2278.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.  
XX WO2004035732-A2.  
XX  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 29-AUG-2002; 2002US-0410946P.  
PR 29-AUG-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486466P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3509; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC anti-inflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 153 AA;  
XX  
Query Match 72.0%; Score 18; DB 8; Length 153;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXX 9  
Db 79 CTATTTTTC 87  
RESULT 20  
ADN23734  
ID ADN23734 standard; protein; 155 AA.  
XX  
XX ADN23734;  
AC  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Bacterial polypeptide #6387.  
XX  
XX Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
OS Bacteria.  
XX  
PN US2003233675-A1.  
XX  
PD 18-DEC-2003.  
XX  
PF 20-FEB-2003; 2003US-00369493.  
XX  
PR 21-FEB-2002; 2002US-0360039P.  
XX  
XX (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX  
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;



XX DR WPI; 2004-061375/06.  
XX PT New recombinant DNA construct comprising a promoter positioned to provide  
PT expression of a polynucleotide encoding a polypeptide from a  
XX PT microbial source, useful for producing plants with improved properties.  
XX PS Claim 1, SEQ ID NO 6387; 122pp; English.  
XX CC The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomanan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
SQ Sequence 155 AA;  
Query Match 72.0%; Score 18; DB 8; Length 155;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 CXXXXXXC 9  
DB 92 CTTASSSSC 100  
RESULT 21  
ABO83103  
XX ID ABO83103 standard; protein; 156 AA.  
XX AC ABO83103;  
XX DT 29-JUL-2004 (first entry)  
XX DE Pseudomonas aeruginosa polypeptide #15278.  
XX DE Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX KM Pseudomonas aeruginosa.  
XX OS US651795-B1.  
XX PN 22-APR-2003.  
XX PD 18-FEB-1999; 99US-00252991.  
XX PF 18-FEB-1999; 98US-0074788P.  
XX PR 27-JUL-1998; 98US-0094190P.  
XX XX  
XX PA (GENO-) GENOME THERAPEUTICS CORP.  
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX WPI; 2003-615309/58.

DR N-PSDB; ABD16674.  
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX PT pathological conditions resulting from bacterial infection.  
XX PS Disclosure; SEQ ID NO 31849; 455pp; English.  
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
SQ Sequence 156 AA;  
Query Match 72.0%; Score 18; DB 7; Length 156;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 CXXXXXXC 9  
DB 46 CASAAATAC 54  
RESULT 22  
ADP30759  
XX ID ADP30759 standard; protein; 156 AA.  
XX AC ADP30759;  
XX DT 12-AUG-2004 (first entry)  
XX DE Human secreted protein SEQ ID #1526.  
XX DE Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX KW cancer; inflammatory; immune; human secreted protein.  
XX OS Homo sapiens.  
XX PN WO2004035732-A2.  
XX PD 29-APR-2004.  
XX PF 28-AUG-2003; 2003WO-US026780.  
XX PR 29-AUG-2002; 2002US-0406576P.  
XX PR 29-AUG-2002; 2002US-0406579P.  
XX PR 29-AUG-2002; 2002US-0406585P.  
XX PR 29-AUG-2002; 2002US-0406588P.  
XX PR 29-AUG-2002; 2002US-0406608P.  
XX PR 29-AUG-2002; 2002US-0406611P.  
XX PR 29-AUG-2002; 2002US-0406612P.  
XX PR 29-AUG-2002; 2002US-0406616P.  
XX PR 29-AUG-2002; 2002US-0406640P.  
XX PR 29-AUG-2002; 2002US-0406642P.  
XX PR 29-AUG-2002; 2002US-0406646P.  
XX PR 29-AUG-2002; 2002US-0406653P.  
XX PR 29-AUG-2002; 2002US-0406655P.  
XX PR 29-AUG-2002; 2002US-0406666P.  
XX PR 17-SEP-2002; 2002US-0410946P.  
XX PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-041101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476699P.  
PR 09-JUN-2003; 2003US-047661P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-048646P.  
PR 14-JUL-2003; 2003US-048648P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-048696P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,  
XX  
XX WPI; 2004-348438/32.  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2757; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present

CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 156 AA;  
Query Match 72.0%; Score 18; DB 8; Length 156;  
Best Local Similarity 22.2%; Pred. No. 3.5e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 CXXXXXXKC 9  
Db 47 CTATTTTAC 55  
RESULT 23  
ADP31286  
ID ADP31286 standard; protein; 165 AA.  
XX  
AC ADP31286;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2053.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411010P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476611P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 14-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang WM, Kochakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 3284; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
XX  
SQ Sequence 165 AA;  
XX  
XX  
Query Match 72.0%; Score 18; DB 8; Length 165;  
Best Local Similarity 22.2%; Pred. No. 3.6e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
OY 1 CXXXXXXC 9  
Db 107 CTTTAATC 115  
XX  
XX  
RESULT 24  
ID ADP31174 standard; protein; 165 AA.  
XX  
XX ADP31174;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
XX

DE Human secreted protein SEQ ID #1941.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
XX  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410944P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410966P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411101P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 18-APR-2003; 2003US-0463700P.  
XX 18-APR-2003; 2003US-0463708P.  
XX 18-APR-2003; 2003US-0463716P.  
XX 18-APR-2003; 2003US-0463732P.  
XX 02-MAY-2003; 2003US-0467199P.  
XX 02-MAY-2003; 2003US-0467201P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 02-MAY-2003; 2003US-0467230P.  
XX 19-MAY-2003; 2003US-0471306P.  
XX 19-MAY-2003; 2003US-0471336P.  
XX 22-MAY-2003; 2003US-0472420P.  
XX 22-MAY-2003; 2003US-0472430P.  
XX 09-JUN-2003; 2003US-0476609P.  
XX 09-JUN-2003; 2003US-0476611P.  
XX 08-JUL-2003; 2003US-0485218P.  
XX 08-JUL-2003; 2003US-0485223P.  
XX 08-JUL-2003; 2003US-0485224P.  
XX 08-JUL-2003; 2003US-0485325P.  
XX

PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halsebeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3172; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.  
XX  
XX  
SQ Sequence 165 AA;

Query Match 72.0%; Score 18; DB 8; Length 165;  
Best Local Similarity 22.2%; Pred. No. 3.6e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
DB 107 CTTTAATC 115

RESULT 25  
ABO83337  
ID ABO83337 standard; protein; 169 AA.  
XX  
XX ABO83337;  
AC  
XX  
XX 29-JUL-2004 (first entry)  
DT  
XX  
XX Pseudomonas aeruginosa polypeptide #15512.  
DE  
XX  
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX  
XX Pseudomonas aeruginosa.  
OS  
XX US6551795-B1.  
PN  
XX  
XX 22-APR-2003.  
PD  
XX  
XX 18-FEB-1999; 99US-00252991.  
PF  
XX  
XX 18-FEB-1998; 98US-0074788P.  
PR 27-JUL-1998; 98US-0094190P.  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
PA  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
PI WPI; 2003-615309/58.  
DR N-PSDB; ABD16908.  
XX

PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.

PS Disclosure; SEQ ID NO 32083; 455bp; English.

XX  
XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)

XX  
SQ Sequence 169 AA;

Query Match 72.0%; Score 18; DB 7; Length 169;  
Best Local Similarity 22.2%; Pred. No. 3.6e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
DB 84 CTTTSAAC 92

RESULT 26  
ADP30801  
ID ADP30801 standard; protein; 174 AA.  
XX  
XX ADP30801;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
DT  
XX  
XX Human secreted protein SEQ ID #1568.  
DE  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2004035732-A2.  
PN  
XX  
XX 29-APR-2004.  
PD  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.

[illegible]

```

SQ      Sequence 174 AA;
Query Match      72.0%; Score 18; DB 8; Length 174;
Best Local Similarity 22.2%; Pred. No. 3.6e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0
OY      1 CXXXXXXXXC 9
Db      122 CTAATTTTC 130

RESULT 27
ID      ABO73977 standard; protein; 193 AA.
XX      ABO73977
AC      ABO73977;
XX      29-JUL-2004 (first entry)
DE      Pseudomonas aeruginosa polypeptide #6152.
XX      Pseudomonas aeruginosa infection; antibacterial.
XX      Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX      Pseudomonas aeruginosa.
OS      US6551795-B1.
XX      22-APR-2003.
XX      18-FEB-1999; 99US-00252991.
XX      18-FEB-1998; 98US-0074788P.
XX      27-JUL-1998; 98US-0094190P.
XX      (GENO-) GENOME THERAPEUTICS CORP.
XX      Rubenfeld MJ, Nolling J, Deloughery C, Bush D;
XX      WPI: 2003-615309/58.
XX      N-PSDB: ABD07548.
XX      Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX      useful as molecular targets for diagnostics, prophylaxis and treatment of
XX      pathological conditions resulting from bacterial infection.
XX      Disclosure; SEQ ID NO 22723; 455pp; English.
XX      The invention relates to Pseudomonas aeruginosa polypeptides and the
XX      polynucleotides encoding them. The sequences are useful in diagnosis and
XX      therapy of pathological conditions, as molecular targets for diagnostics,
XX      prophylaxis and treatment of pathological conditions resulting from a
XX      bacterial infection, for evaluating a compound, such as a polypeptide,
XX      for the ability to bind a P. aeruginosa nucleic acid, as components of
XX      effective antibacterial targets, as targets for antibacterial drugs,
XX      including anti-P. aeruginosa drugs, as templates for recombinant
XX      production of P. aeruginosa-derived peptides or polypeptides, as target
XX      components for diagnosis and/or treatment of P. aeruginosa-caused
XX      infection, and in detection of P. aeruginosa sequences or other sequences
XX      of Pseudomonas species using biotech technology. Sequences ABO67826-
XX      ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX      sequence data for this patent did not form part of the printed
XX      specification but was obtained in electronic format from USPRO at
XX      seqdata.uspro.gov/sequence.html
XX      Sequence 193 AA;
Query Match      72.0%; Score 18; DB 7; Length 193;
Best Local Similarity 22.2%; Pred. No. 3.7e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY      1 CXXXXXXXXC 9

```

Db	61	CTSTSATC	69
RESULT	28		
ID	ADP30493		
XX	ADP30493	standard; protein; 198 AA.	
AC	ADP30493;		
DT	12-AUG-2004	(first entry)	
XX			
DE	Human secreted protein SEQ ID #1260.		
XX			
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;		
KW	Cancer; inflammatory; immune; human secreted protein.		
OS	Homo sapiens.		
XX			
PN	WO2004035732-A2.		
XX			
PD	29-APR-2004.		
PF	28-AUG-2003;	2003WO-US026780.	
XX			
PR	29-AUG-2002;	2002US-0406576P.	
PR	29-AUG-2002;	2002US-0406579P.	
PR	29-AUG-2002;	2002US-0406585P.	
PR	29-AUG-2002;	2002US-0406588P.	
PR	29-AUG-2002;	2002US-0406608P.	
PR	29-AUG-2002;	2002US-0406611P.	
PR	29-AUG-2002;	2002US-0406612P.	
PR	29-AUG-2002;	2002US-0406616P.	
PR	29-AUG-2002;	2002US-0406640P.	
PR	29-AUG-2002;	2002US-0406642P.	
PR	29-AUG-2002;	2002US-0406646P.	
PR	29-AUG-2002;	2002US-0406653P.	
PR	29-AUG-2002;	2002US-0406655P.	
PR	29-AUG-2002;	2002US-0406666P.	
PR	17-SEP-2002;	2002US-0410946P.	
PR	17-SEP-2002;	2002US-0410947P.	
PR	17-SEP-2002;	2002US-0410949P.	
PR	17-SEP-2002;	2002US-0410953P.	
PR	17-SEP-2002;	2002US-0410957P.	
PR	17-SEP-2002;	2002US-0410958P.	
PR	17-SEP-2002;	2002US-0410959P.	
PR	17-SEP-2002;	2002US-0410960P.	
PR	17-SEP-2002;	2002US-0410961P.	
PR	17-SEP-2002;	2002US-0410962P.	
PR	17-SEP-2002;	2002US-0411019P.	
PR	17-SEP-2002;	2002US-0411022P.	
PR	17-SEP-2002;	2002US-0411023P.	
PR	17-SEP-2002;	2002US-0411024P.	
PR	17-SEP-2002;	2002US-0411032P.	
PR	17-SEP-2002;	2002US-0411035P.	
PR	17-SEP-2002;	2002US-0411037P.	
PR	17-SEP-2002;	2002US-0411041P.	
PR	17-SEP-2002;	2002US-0411045P.	
PR	17-SEP-2002;	2002US-0411046P.	
PR	17-SEP-2002;	2002US-0411048P.	
PR	17-SEP-2002;	2002US-0411052P.	
PR	17-SEP-2002;	2002US-0411055P.	
PR	17-SEP-2002;	2002US-0411073P.	
PR	17-SEP-2002;	2002US-0411082P.	
PR	17-SEP-2002;	2002US-0411101P.	
PR	17-SEP-2002;	2002US-0411111P.	
PR	18-APR-2003;	2003US-0463700P.	
PR	18-APR-2003;	2003US-0463708P.	
PR	18-APR-2003;	2003US-0463716P.	
PR	18-APR-2003;	2003US-0463719P.	
PR	02-MAY-2003;	2003US-0467199P.	
PR	02-MAY-2003;	2003US-0467201P.	
PR	02-MAY-2003;	2003US-0467203P.	

PR	02-MAY-2003;	2003US-0467230P.
PR	19-MAY-2003;	2003US-0471306P.
PR	19-MAY-2003;	2003US-0471336P.
PR	22-MAY-2003;	2003US-0472420P.
PR	09-JUN-2003;	2003US-0476609P.
PR	09-JUN-2003;	2003US-0476641P.
PR	08-JUL-2003;	2003US-0485218P.
PR	08-JUL-2003;	2003US-0485223P.
PR	08-JUL-2003;	2003US-0485224P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486446P.
PR	14-JUL-2003;	2003US-0486480P.
PR	15-JUL-2003;	2003US-0486891P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.
XX		
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX		
PI	Williams LT, Chu K, Lee E, Heetir K, Beaurang PA, Behrens D;	
PI	Halenbeck RF, Huang MW, Kothakota S, Halshan L, Dimmema T;	
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,	
XX		
DR	WPI; 2004-348438/32.	
XX		
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
PS	Claim 1; SEQ ID NO 2491; 428bp; English.	
XX		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytosolic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
CC	available on WIPOMEB and is not in the specification.	
XX		
SEQ	Sequence 198 AA:	
QY	1 CXXXXXXKC 9	
Db	91 CATTTTTC 99	
RESULT 29		
ADP30492		
ID	ADP30492 standard; protein; 198 AA.	
XX		
AC	ADP30492;	
XX		
DT	12-AUG-2004 (first entry)	
XX		
DE	Human secreted protein SEQ ID #1259.	
XX		
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;	
XX	cancer; inflammatory; immune; human secreted protein.	
XX		
OS	Homo sapiens.	
XX		
PN	WO2004035732-A2.	
XX		
PD	29-APR-2004.	
XX		

PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486911P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493340P.  
PR 08-AUG-2003; 2003US-0493533P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;  
PI Pierre K, Wang Y, Wong JCP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1, SEQ ID NO 2490; 428pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC anti-inflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
XX available on WIPOMEB and is not in the specification.  
SQ Sequence 198 AA;  
Query Match 72.0%; Score 18; DB 8; Length 198;  
Best Local Similarity 22.2%; Pred. No. 3.8e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 CXXXXXXC 9  
Db 91 CATTYTTTC 99  
RESULT 30  
ADP30477  
ID ADP30477 standard; protein, 198 AA.  
XX  
AC ADP30477;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1244.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410969P.  
PR 17-SEP-2002; 2002US-0411012P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467206P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485244P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493331P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Heestir K, Beaurang PA, Behrens D;  
PI Halerbeck RF, Huang WM, Kochakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 2475; 428pp; English.  
XX  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytosstatic,  
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present  
XX sequence represents a human secreted protein. The present sequence is  
XX available on WIPOWEB and is not in the specification.  
XX  
XX Sequence 198 AA:

Query Match 72.0%; Score 18; DB 8; Length 198;  
Best Local Similarity 22.2%; Pred: No. 3.8e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXX 9  
Db 54 CATATTAC 62  
RESULT 31  
ADP30481  
ID ADP30481 standard; protein; 198 AA.  
XX  
AC ADP30481;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #1248.  
DE  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX  
XX MO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410944P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411101P.  
XX 17-SEP-2002; 2002US-0411111P.







CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.

XX  
SQ Sequence 228 AA;

Query Match Best Local Similarity 72.0%; Score 18; DB 8; Length 228;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
DB 183 CTTTAATC 191

RESULT 34  
ABO80501  
ID ABO80501 standard; protein; 229 AA.

XX ABO80501;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #12676.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US651795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX MPI: 2003-615309/58.

XX N-PSDB; ABD14072.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 29247; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using bioclip technology. Sequences ABO67826-  
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html

XX Sequence 229 AA;

Query Match 72.0%; Score 18; DB 7; Length 229;

Best Local Similarity 22.2%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
DB 155 CATTATAC 163

RESULT 35

ABO77555  
ID ABO77555 standard; protein; 233 AA.

XX ABO77555;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #9710.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US651795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX MPI: 2003-615309/58.

XX N-PSDB; ABD11126.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 26301; 455bp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX polynucleotides encoding them. The sequences are useful in diagnosis and  
XX therapy of pathological conditions, as molecular targets for diagnostics,  
XX prophylaxis and treatment of pathological conditions resulting from a  
XX bacterial infection, for evaluating a compound, such as a polypeptide,  
XX for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX effective antibacterial targets, as targets for antibacterial drugs,  
XX including anti-P. aeruginosa drugs, as templates for recombinant  
XX production of P. aeruginosa-derived peptides or polypeptides, as target  
XX components for diagnosis and/or treatment of P. aeruginosa-caused  
XX infection, and in detection of P. aeruginosa sequences or other sequences  
XX of Pseudomonas species using bioclip technology. Sequences ABO67826-  
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX sequence data for this patent did not form part of the printed  
XX specification but was obtained in electronic format from USPTO at  
XX seqdata.uspto.gov/sequence.html

XX Sequence 233 AA;

Query Match 72.0%; Score 18; DB 7; Length 233;

Best Local Similarity 22.2%; Pred. No. 3.9e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
DB 43 CATTATAC 51

RESULT 36

```
ADP31468
ID ADP31468 standard; protein; 234 AA.
XX
AC ADP31468;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2235.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
XX
PR 29-AUG-2002; 2002US-0406585P.
XX
PR 29-AUG-2002; 2002US-0406588P.
XX
PR 29-AUG-2002; 2002US-0406608P.
XX
PR 29-AUG-2002; 2002US-0406611P.
XX
PR 29-AUG-2002; 2002US-0406612P.
XX
PR 29-AUG-2002; 2002US-0406616P.
XX
PR 29-AUG-2002; 2002US-0406640P.
XX
PR 29-AUG-2002; 2002US-0406642P.
XX
PR 29-AUG-2002; 2002US-0406646P.
XX
PR 29-AUG-2002; 2002US-0406653P.
XX
PR 29-AUG-2002; 2002US-0406655P.
XX
PR 29-AUG-2002; 2002US-0406666P.
XX
PR 17-SEP-2002; 2002US-0410946P.
XX
PR 17-SEP-2002; 2002US-0410947P.
XX
PR 17-SEP-2002; 2002US-0410948P.
XX
PR 17-SEP-2002; 2002US-0410949P.
XX
PR 17-SEP-2002; 2002US-0410953P.
XX
PR 17-SEP-2002; 2002US-0410957P.
XX
PR 17-SEP-2002; 2002US-0410958P.
XX
PR 17-SEP-2002; 2002US-0410959P.
XX
PR 17-SEP-2002; 2002US-0410960P.
XX
PR 17-SEP-2002; 2002US-0410961P.
XX
PR 17-SEP-2002; 2002US-0410962P.
XX
PR 17-SEP-2002; 2002US-0411019P.
XX
PR 17-SEP-2002; 2002US-0411022P.
XX
PR 17-SEP-2002; 2002US-0411023P.
XX
PR 17-SEP-2002; 2002US-0411024P.
XX
PR 17-SEP-2002; 2002US-0411032P.
XX
PR 17-SEP-2002; 2002US-0411035P.
XX
PR 17-SEP-2002; 2002US-0411037P.
XX
PR 17-SEP-2002; 2002US-0411041P.
XX
PR 17-SEP-2002; 2002US-0411045P.
XX
PR 17-SEP-2002; 2002US-0411046P.
XX
PR 17-SEP-2002; 2002US-0411048P.
XX
PR 17-SEP-2002; 2002US-0411052P.
XX
PR 17-SEP-2002; 2002US-0411055P.
XX
PR 17-SEP-2002; 2002US-0411073P.
XX
PR 17-SEP-2002; 2002US-0411082P.
XX
PR 17-SEP-2002; 2002US-0411101P.
XX
PR 17-SEP-2002; 2002US-0411111P.
XX
PR 18-APR-2003; 2003US-0463700P.
XX
PR 18-APR-2003; 2003US-0463708P.
XX
PR 18-APR-2003; 2003US-0463716P.
XX
PR 18-APR-2003; 2003US-0463732P.
XX
PR 02-MAY-2003; 2003US-0467199P.
XX
PR 02-MAY-2003; 2003US-0467201P.
XX
PR 02-MAY-2003; 2003US-0467203P.
XX
PR 02-MAY-2003; 2003US-0467230P.
XX
PR 19-MAY-2003; 2003US-0471306P.
XX
PR 19-MAY-2003; 2003US-0471336P.
XX
PR 22-MAY-2003; 2003US-0472420P.
XX
PR 22-MAY-2003; 2003US-0472430P.
XX
PR 09-JUN-2003; 2003US-0476609P.
XX
PR 09-JUN-2003; 2003US-0476641P.
XX
PR 08-JUL-2003; 2003US-0485218P.
XX
PR 08-JUL-2003; 2003US-0485223P.
XX
PR 08-JUL-2003; 2003US-0485224P.
XX
PR 08-JUL-2003; 2003US-0485325P.
XX
PR 14-JUL-2003; 2003US-0486446P.
XX
PR 14-JUL-2003; 2003US-0486480P.
XX
PR 15-JUL-2003; 2003US-0486891P.
XX
PR 15-JUL-2003; 2003US-0486960P.
XX
PR 08-AUG-2003; 2003US-0493341P.
XX
PR 08-AUG-2003; 2003US-0493370P.
XX
PR 08-AUG-2003; 2003US-0493573P.
XX
PR 08-AUG-2003; 2003US-0493577P.
XX
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
DR WPI; 2004-348438/32.
XX
PT New nucleic acid molecule for diagnosing, preventing or treating diseases
XX
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX
PT genetic, bacterial and viral diseases.
XX
PS Claim 1; SEQ ID NO 3466; 428bp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
XX
CC encoding a polypeptide which is believed to be cytostatic,
XX
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX
CC composition and methods are useful for diagnosing, preventing and
XX
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX
CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX
CC sequence represents a human secreted protein. The present sequence is
XX
CC available on WIPOWEB and is not in the specification.
XX
SQ Sequence 234 AA;
XX
Query Match 72.0%; Score 18; DB 8; Length 234;
Best Local Similarity 22.2%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXXC 9
Db 213 CTTATTAC 221
RESULT 37
ADP31616
ID ADP31616 standard; protein; 242 AA.
XX
AC ADP31616;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #2383.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
PN WO2004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
```

PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406642P.  
 PR 29-AUG-2002; 2002US-0406646P.  
 PR 29-AUG-2002; 2002US-0406653P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 29-AUG-2002; 2002US-0406666P.  
 PR 17-SEP-2002; 2002US-0410946P.  
 PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410948P.  
 PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410958P.  
 PR 17-SEP-2002; 2002US-0410959P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410961P.  
 PR 17-SEP-2002; 2002US-0410962P.  
 PR 17-SEP-2002; 2002US-0411019P.  
 PR 17-SEP-2002; 2002US-0411022P.  
 PR 17-SEP-2002; 2002US-0411023P.  
 PR 17-SEP-2002; 2002US-0411024P.  
 PR 17-SEP-2002; 2002US-0411032P.  
 PR 17-SEP-2002; 2002US-0411035P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411041P.  
 PR 17-SEP-2002; 2002US-0411045P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411101P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471366P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485255P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 PR XX  
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX  
 PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
 PI Halebek R, Huang MM, Kochakota S, Haishan L, Linnemann T,  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,  
 XX

DR WPI; 2004-348438/32.  
 XX  
 XX New nucleic acid molecule for diagnosis, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX  
 XX Claim 1; SEQ ID NO 3614; 428bp; English.  
 XX  
 CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic.  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMEB and is not in the specification.  
 XX  
 SQ Sequence 242 AA;

Query Match 72.0%; Score 18; DB 8; Length 242;  
 Best Local Similarity 22.2%; Pred. No. 46+02;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 Db 148 CTAATAAC 156

RESULT 38  
 AAU75930  
 ID AAU75930 standard; protein; 246 AA.  
 XX

AC AAU75930;

DT 08-MAY-2002 (first entry)

DE Androgen receptor ligand binding domain (AR-LBD).

XX  
 KW Androgen; receptor; ligand binding domain; human;  
 KW androgen receptor modulator; SAR; hormone-dependent tumour;  
 KW prostate cancer; hirsutism; acne; seborrhea; Alzheimer's disease;  
 KW androgenic alopecia; hypogonadism; hyperlipidosis; prostate cancer;  
 KW benign prostate hyperplasia; adenoma; tumour; pancreatic cancer;  
 KW vascular endothelial growth factor; VEGF; antiangiogenic agent;  
 KW osteoporosis; spermatogenesis; libido; cachexia; endometriosis;  
 KW polycystic ovary syndrome; anorexia; male menopause;  
 KW male hormone replacement; female sexual dysfunction; gynaecomastia;  
 KW male sexual dysfunction; muscular atrophy; muscle wasting;  
 KW protein co-ordinate data.  
 XX

OS Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

WO200200617-A2.

```

XX 03-JAN-2002.
PD
XX
XX 20-JUN-2001; 2001WO-US019665.
PR
XX
XX 28-JUN-2000; 2000US-0214392P.
PR 19-SEP-2000; 2000US-0233519P.
PR 18-APR-2001; 2001US-0284438P.
PR 18-APR-2001; 2001US-0284617P.
PR 18-APR-2001; 2001US-0284730P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
PA
XX Salvati ME, Gottardis MM, Kryetex SR, Altar RM, Sack JS;
PI
XX WPI; 2002-164433/21.
DR
XX
XX PT New selective androgen receptor modulator for treating tumors, comprises
XX PT antagonist activity in a hormone-dependent tumor and no or agonist
XX PT activity against other, non-tumor tissues containing the receptor.
XX
XX PS Disclosure; Page 19; 140pp; English.
XX
XX CC The invention describes a selective androgen receptor modulator (SARM)
XX CC (I) which exhibits antagonist activity in a hormone-dependent tumor
XX CC while exhibiting no activity or agonist activity against other non-tumor
XX CC tissues containing the androgen receptor. (II) is useful for inhibiting
XX CC the growth of hormone-dependent tumor cells (prostate tumor cells) in a
XX CC patient. (I) is also useful for treating a condition selected from
XX CC hirsutism, acne, seborrhea, Alzheimer's disease, androgenic alopecia,
XX CC hypogonadism, hyperlipidosis, benign prostate hypertrophy, adenomas or
XX CC neoplasms of the prostate, treatment of benign or malignant tumor cells
XX CC containing the androgen receptor, pancreatic cancer, modulation of
XX CC vascular endothelial growth factor (VEGF) expression for use as
XX CC antiangiogenic agents, osteoporosis, suppressing spermatogenesis, libido,
XX CC cachexia, endometriosis, polycystic ovary syndrome, anorexia, androgen
XX CC dependent age-related diseases and conditions, male menopause, male
XX CC hormone replacement, male and female sexual dysfunction, and inhibition
XX CC of muscular atrophy in ambulatory patients. (II) is also useful for
XX CC treating muscle wasting and gynecomastia. This sequence represents the
XX CC androgen receptor ligand binding domain (AR-LBD), described in the method
XX CC of the invention
XX
XX SQ Sequence 246 AA;

```

```

Query Match 72.0%; Score 18; DB 5; Length 246;
Best Local Similarity 22.2%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 1 CXXXXXXC 9
Db 173 CAAAAAASC 181

```

```

RESULT 39
ABO74848
ID ABO74848 standard; protein; 246 AA.
XX
XX ABO74848;
AC
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Pseudomonas aeruginosa polypeptide #7023.
DE
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
KM
XX
XX Pseudomonas aeruginosa.
OS
XX
XX US6551795-B1.
PN
XX
XX 22-APR-2003.
PD
XX
XX 18-FEB-1999; 99US-00252991.
PF

```

```

XX 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI
XX WPI; 2003-615309/58.
DR
XX N-PSDB; ABD08419.
XX
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX
XX PS Disclosure; SEQ ID NO 23594; 455pp; English.
XX
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-P. aeruginosa drugs, as templates for recombinant
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused
XX CC infection, and in detection of P. aeruginosa sequences or other sequences
XX CC of Pseudomonas species using biochip technology. Sequences ABO67826-
XX CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format from USPTO at
XX CC segdata.uspto.gov/sequence.html
XX
XX SQ Sequence 246 AA;

```

```

Query Match 72.0%; Score 18; DB 7; Length 246;
Best Local Similarity 22.2%; Pred. No. 4e+02;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 1 CXXXXXXC 9
Db 202 CSTTSATC 210

```

```

RESULT 40
ADP30754
ID ADP30754 standard; protein; 249 AA.
XX
XX ADP30754;
AC
XX
XX 12-AUG-2004 (first entry)
DT
XX
XX Human secreted protein SEQ ID #1521.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KM cancer; inflammatory; immune; human secreted protein.
XX
XX Homo sapiens.
OS
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
PR 29-AUG-2002; 2002US-0406579P.
PR 29-AUG-2002; 2002US-0406585P.
PR 29-AUG-2002; 2002US-0406588P.
PR 29-AUG-2002; 2002US-0406608P.
PR 29-AUG-2002; 2002US-0406611P.
PR 29-AUG-2002; 2002US-0406612P.
PR

```

PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485255P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX MPI; 2004-348438/32.  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

XX  
XX Claim 1; SEQ ID NO 2752; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic. The  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
XX Sequence 249 AA;  
SQ  
Query Match 72.0%; Score 18; DB 8; Length 249;  
Best Local Similarity 22.2%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXC 9  
Db 238 CTRAAAAAC 246  
RESULT 41  
ID ADP31396 standard; protein; 254 AA.  
XX  
AC ADP31396;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2163.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486919P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
(FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX PA Williams LR, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3394; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 254 AA;

Query Match 72.0%; Score 18; DB 8; Length 254;  
Best Local Similarity 22.2%; Pred. No. 4e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 CXXXXXXXXC 9  
DB 77 CTATAAAC 85

RESULT 42  
ADP30846  
ID ADP30846 standard; protein; 279 AA.  
XX  
XX ADP30846;  
XX  
XX 12-AUG-2004 (first entry)  
DT  
XX  
DE Human secreted protein SEQ ID #1613.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
XX  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411101P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 18-APR-2003; 2003US-0463700P.  
XX 18-APR-2003; 2003US-0463708P.  
XX 18-APR-2003; 2003US-0463716P.  
XX 18-APR-2003; 2003US-0463732P.  
XX 02-MAY-2003; 2003US-0467199P.  
XX 02-MAY-2003; 2003US-0467201P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 02-MAY-2003; 2003US-0467230P.  
XX 19-MAY-2003; 2003US-0471306P.



PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476611P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486460P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 2844; 428bp; English.  
XX  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
XX  
SQ Sequence 279 AA;  
  
Query Match 72.0%; Score 18; DB 8; Length 279;  
Best Local Similarity 22.2%; Pred. No. 4.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXX 9  
DB 247 CTAATTAAC 255  
  
RESULT 43  
ADP31523  
ID ADP31523 standard; protein; 279 AA.  
XX  
XX  
AC ADP31523;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2290.  
XX  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX  
PN MO2004035732-A2.  
XX  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX

PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476611P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0486446P.  
PR 08-JUL-2003; 2003US-0486460P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3521; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
XX available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 279 AA;  
Query Match 72.0%; Score 18; DB 8; Length 279;  
Best Local Similarity 22.2%; Pred. No. 4.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 CXXXXXXXC 9  
DB 66 CATTAAATC 74  
RESULT 44  
ADP30844  
ID ADP30844 standard; protein; 279 AA.  
XX  
XX ADP30844;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #1611.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX  
XX 29-APR-2004.  
XX  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 29-AUG-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485233P.  
PR 08-JUL-2003; 2003US-0485242P.  
PR 08-JUL-2003; 2003US-0485252P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2842; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
XX available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 279 AA;  
Query Match 72.0%; Score 18; DB 8; Length 279;

Best Local Similarity 22.2%; Pred. No. 4.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 247 CTATTTAAC 255

## RESULT 45

AB48050  
ID AB48050 standard; protein, 281 AA.

AC AAB48050;

DT 19-MAR-2001 (first entry)

DE Signal transduction protein.

KW Zea mays; maize; signal transduction protein; phytohormone; ethylene;

XX auxin; cytokinin; gibberellin; immunogen.

OS Zea mays.

PN WO20070059-A2.

PD 23-NOV-2000.

PF 28-APR-2000; 2000WO-US011687.

PR 14-MAY-1999; 99US-0134292P.

PR 08-JUL-1999; 99US-0142996P.

PA (PION-) PIONEER HI-BRED INT INC.

PI Helencjaris TG;

DR WPI; 2001-031929/04.

DR N-PSDB; AAC84271.

PT New signal transduction nucleic acids and encoded proteins useful for  
PT regulating phytohormone expression, including ethylene, auxins,  
PT cytokinins and gibberellin, to provide control of plant response to  
PT environmental stresses.

PS Claim 13; Page 102; 126pp; English.

CC The invention provides Zea mays signal transduction proteins and encoding  
CC nucleotide sequences. The nucleic acids are useful for regulating  
CC expression of phytohormones, including ethylene, auxins, cytokinins, and  
CC gibberellin, to effect developmental changes in plants and provide  
CC control of plant response to environmental stresses. They may also be  
CC used as probes or amplification primers in the detection, quantitation or  
CC isolation of gene transcripts, for detecting mutations in the gene, for  
CC monitoring upregulation of expression or changes in enzyme activity in  
CC screening assays of compounds, for detection of any number of allelic  
CC variants, or for site-directed mutagenesis in eukaryotic cells. They may  
CC further be used for recombinant expression of their encoded polypeptides,  
CC as immunogens in the preparation or screening of antibodies, and in sense  
CC or antisense suppression of genes in a host cell, tissue or plant. The  
CC proteins may be used in assays for enzyme agonists or antagonists, as  
CC immunogens or antigens to obtain antibodies specifically immunoreactive  
CC with the proteins. The present sequence represents a signal transduction  
CC protein of the invention

XX Sequence 281 AA;

Query Match 72.0%; Score 18; DB 4; Length 281;

Best Local Similarity 22.2%; Pred. No. 4.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 11 CSATTTSSC 19

## RESULT 46

ADM43729  
ID ADM43729 standard; protein, 281 AA.

AC ADM43729;

DT 24-MAR-2005 (first entry)

DE Corn DREB-type transcription factor, seq id 2.

KW Drought resistance; crop improvement; cold tolerance;

XX transcription factor; DREB.

OS Zea mays.

PN CN1472222-A.

PD 04-FEB-2004.

PF 29-JUL-2002; 2002CN-00125372.

PR 29-JUL-2002; 2002CN-00125372.

PA (UYOI ) UNIV QINGHUA.

PI Liu Q, Qin F, Zhao J;

DR WPI; 2004-317379/30.

DR N-PSDB; ADM43728.

PT DREB transcription factor of corn and its encoding genes and use.

PS Disclosure; SEQ ID NO 2; 18pp; Chinese.

CC The present invention discloses 2 corn DREB-type transcription factors,  
CC madREB1 and madREB2, which are important in conditions of cold and  
CC drought. Also disclosed are their coding genes and their application in  
CC culture the cold and drought resistant plants. The current sequence  
CC represents a DREB-type transcription factor.

XX Sequence 281 AA;

Query Match 72.0%; Score 18; DB 8; Length 281;

Best Local Similarity 22.2%; Pred. No. 4.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 11 CSATTTSSC 19

## RESULT 47

AB079000  
ID AB079000 standard; protein, 286 AA.

AC AB079000;

DT 29-JUL-2004 (first entry)

DE Pseudomonas aeruginosa polypeptide #11175.

KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

PN US6551795-B1.

PD 22-APR-2003.

PF 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.  
PR 27-JUL-1998; 98US-0094190P.  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX  
XX  
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX WPI: 2003-615109/58.  
DR N-PSDB; ABD12571.  
XX  
XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
PS Disclosure; SEQ ID NO 27746; 455bp; English.  
XX  
XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-*P. aeruginosa* drugs, as templates for recombinant  
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
CC of *Pseudomonas* species using bioclip technology. Sequences AB067826-  
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)  
SQ Sequence 286 AA;  
  
Query Match 72.0%; Score 18; DB 7; Length 286;  
Best local Similarity 22.2%; Pred. No. 4.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXC 9  
Db 131 CCGSNAATAC 139  
  
RESULT 48  
ADP31192  
ID ADP31192 standard; protein; 297 AA.  
XX  
AC ADP31192;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #1959.  
DE  
XX Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
OS  
XX WO2004035732-A2.  
PN  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PA  
XX  
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kotchakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI: 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX

PS Claim 1; SEQ ID NO 3190; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 297 AA;  
XX  
Query Match 72.0%; Score 18; DB 8; Length 297;  
Best Local Similarity 22.2%; Pred. No. 4.2e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
OY 1 CXXXXXXC 9  
Db 39 CAATTTAC 47  
XX  
RESULT 49  
ADP31454  
ID ADP31454 standard; protein; 312 AA.  
XX  
XX ADP31454;  
AC  
XX 12-AUG-2004 (first entry)  
DT  
XX Human secreted protein SEQ ID #2221.  
DE  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; Inflammatory; Immune; human secreted protein.  
XX  
XX Homo sapiens.  
OS  
XX  
XX MO2004035732-AA.  
PN  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003MO-US026780.  
PF  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR  
XX 29-AUG-2002; 2002US-0406579P.  
PI  
XX 29-AUG-2002; 2002US-0406585P.  
PR  
XX 29-AUG-2002; 2002US-0406588P.  
PR  
XX 29-AUG-2002; 2002US-0406608P.  
PR  
XX 29-AUG-2002; 2002US-0406611P.  
PR  
XX 29-AUG-2002; 2002US-0406612P.  
PR  
XX 29-AUG-2002; 2002US-0406616P.  
PR  
XX 29-AUG-2002; 2002US-0406640P.  
PR  
XX 29-AUG-2002; 2002US-0406642P.  
PR  
XX 29-AUG-2002; 2002US-0406646P.  
PR  
XX 29-AUG-2002; 2002US-0406653P.  
PR  
XX 29-AUG-2002; 2002US-0406655P.  
PR  
XX 29-AUG-2002; 2002US-0406666P.  
PR  
XX 17-SEP-2002; 2002US-0410946P.  
PR  
XX 17-SEP-2002; 2002US-0410947P.  
PR  
XX 17-SEP-2002; 2002US-0410948P.  
PR  
XX 17-SEP-2002; 2002US-0410949P.  
PR  
XX 17-SEP-2002; 2002US-0410953P.  
PR  
XX 17-SEP-2002; 2002US-0410957P.  
PR  
XX 17-SEP-2002; 2002US-0410958P.  
PR  
XX 17-SEP-2002; 2002US-0410959P.  
PR  
XX 17-SEP-2002; 2002US-0410960P.  
PR  
XX 17-SEP-2002; 2002US-0410961P.  
PR  
XX 17-SEP-2002; 2002US-0410962P.  
PR  
XX 17-SEP-2002; 2002US-0411019P.  
PR  
XX 17-SEP-2002; 2002US-0411022P.  
PR  
XX 17-SEP-2002; 2002US-0411023P.  
PR  
XX 17-SEP-2002; 2002US-0411024P.  
PR  
XX 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485225P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486860P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PA  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JCP, Wu G, Zhang H;  
DR WPI, 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 3452; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 312 AA;  
XX  
Query Match 72.0%; Score 18; DB 8; Length 312;  
Best Local Similarity 22.2%; Pred. No. 4.2e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
OY 1 CXXXXXXC 9  
Db 7 CAATTTTC 15

RESULT 50  
ADP31615  
ID ADP31615 standard; protein; 317 AA.  
XX  
XX ADP31615;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
DT  
XX  
XX Human secreted protein SEQ ID #2382.  
DE  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
PN WO2004035732-A2.  
XX  
XX  
XX 29-APR-2004.  
PD  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PP  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485249P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PA  
XX  
XX Williams LR, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RP, Huang MM, Kothakota S, Haislan L, Linemann T;  
PI Pierce K, Wang Y, Wong JCB, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3613; 428pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
XX Sequence 317 AA;  
SQ

Query Match 72.0%; Score 18; DB 8; Length 317;  
Best Local Similarity 22.2%; Pred. No. 4.2e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1 CXXXXXXC 9  
DB 145 CATTATATC 153

Search completed: January 4, 2006, 15:56:15  
Job time : 62.5913 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 9.31304 Seconds  
(without alignments)  
92.983 Million cell updates/sec

Title: US-09-932-322-10  
Perfect score: 25  
Sequence: 1 CXXXXXXC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 28316 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 28316

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	72.0	155	2 T28945	hypothetical prote
2	18	72.0	470	2 T45962	hypothetical prote
3	18	72.0	1237	2 A34598	ecdysone-induced p
4	18	72.0	1394	2 B34598	ecdysone-induced p
5	18	72.0	1443	2 S05979	steroid hormone re
6	17	68.0	199	2 T33355	hypothetical prote
7	17	68.0	212	2 S63342	hypothetical prote
8	17	68.0	214	2 C70812	probable lipoq prot
9	17	68.0	244	2 H72532	hypothetical prote
10	17	68.0	248	2 T40415	hypothetical prote
11	17	68.0	259	2 H71353	conserved hypothet
12	17	68.0	281	2 AC0661	probable lipoprote
13	17	68.0	293	2 T48975	xyloglucan endo-tr
14	17	68.0	298	2 T01948	hypothetical prote
15	17	68.0	315	2 A85019	probable BHLH DNA-
16	17	68.0	373	2 AB1224	cobalamin biosynth
17	17	68.0	373	2 AD1577	cobalamin biosynth
18	17	68.0	415	2 T43352	nuclear receptor N
19	17	68.0	507	2 T33024	hypothetical prote
20	17	68.0	530	2 T48004	multifunctional am
21	17	68.0	533	2 A42143	skin-1 - Caenothad
22	17	68.0	543	2 T07739	probable ferriochel
23	17	68.0	575	2 T07953	chromomodulin pre
24	17	68.0	582	2 T07953	lectin-like protei
25	17	68.0	584	2 T07952	lectin-like protei
26	17	68.0	584	2 B25682	homeotic protein B
27	17	68.0	589	2 S41011	hypothetical prote
28	17	68.0	595	2 A42086	CD30 antigen precu
29	17	68.0	600	2 D83286	hypothetical prote

30	17	68.0	617	2 T15408	hypothetical prote
31	17	68.0	624	2 A35962	laccase (EC 1.10.3
32	17	68.0	769	2 A41029	integrin beta-8 ch
33	17	68.0	791	2 A46140	diacylglycerol kin
34	17	68.0	796	2 B46140	diacylglycerol kin
35	17	68.0	799	2 A38308	integrin beta-5 ch
36	17	68.0	934	2 T05201	hypothetical prote
37	17	68.0	938	2 T41932	hypothetical prote
38	17	68.0	938	2 T15881	hypothetical prote
39	17	68.0	1295	1 S50820	surface protein ty
40	17	68.0	2704	2 S09118	G surface protein
41	17	68.0	2718	2 A23475	G surface protein
42	17	68.0	8243	2 T31307	type I fatty acid
43	16	64.0	52	2 G82010	probable lipoprote
44	16	64.0	87	2 H86753	prophage p12 prote
45	16	64.0	91	2 J50036	Clara cell 10K pro
46	16	64.0	96	1 UGMS	uteroglobin precu
47	16	64.0	96	2 A36581	polychlorinated bi
48	16	64.0	98	2 A10334	conserved hypothet
49	16	64.0	103	2 S76421	ferredoxin [2Fe-2S
50	16	64.0	104	2 J00863	hypothetical 11.6K
51	16	64.0	116	2 S09845	hypothetical prote
52	16	64.0	121	2 A44544	hypothetical prote
53	16	64.0	128	2 S74085	amine dehydrogenas
54	16	64.0	138	2 T21792	lutropin beta chal
55	16	64.0	148	2 B72513	hypothetical prote
56	16	64.0	152	2 H97405	hypothetical prote
57	16	64.0	153	2 F75478	conserved hypothet
58	16	64.0	156	2 T31839	hypothetical prote
59	16	64.0	158	2 A61091	lutropin beta chal
60	16	64.0	159	2 I51373	lutropinizing hormon
61	16	64.0	160	1 S17194	endothelin 2 precu
62	16	64.0	166	2 I51242	lutropinizing hormon
63	16	64.0	166	2 S68590	hypothetical prote
64	16	64.0	182	2 T01909	transcription init
65	16	64.0	190	2 S48101	xyloglucan endo-1,
66	16	64.0	195	2 T19151	hypothetical prote
67	16	64.0	207	2 A69446	conserved hypothet
68	16	64.0	209	2 S51480	drought-induced pr
69	16	64.0	238	2 T04166	thumatin-like pro
70	16	64.0	246	2 A63519	enterobactin synth
71	16	64.0	246	2 T20341	hypothetical prote
72	16	64.0	248	2 T21786	hypothetical prote
73	16	64.0	251	2 A55523	hypothetical prote
74	16	64.0	252	2 T04739	hypothetical prote
75	16	64.0	255	2 T38426	lymphocyte activat
76	16	64.0	260	2 T23033	hypothetical prote
77	16	64.0	261	2 C64049	hypothetical prote
78	16	64.0	261	2 S63604	homeobox protein G
79	16	64.0	262	2 S01914	El membrane glycop
80	16	64.0	265	2 H97280	hydroxyethylthiaz
81	16	64.0	266	2 T02807	arsenate reductase
82	16	64.0	266	2 H84409	chlamy biosynthet
83	16	64.0	268	2 S19184	hypothetical prote
84	16	64.0	268	2 H81806	histidine-binding
85	16	64.0	268	2 B81063	amino acid ABC tra
86	16	64.0	269	2 T35647	phosphomethylpyrim
87	16	64.0	275	2 T32813	hypothetical prote
88	16	64.0	276	2 T01177	hypothetical prote
89	16	64.0	276	2 T52349	F-box protein FB12
90	16	64.0	277	2 A46241	interferon respons
91	16	64.0	278	2 T39749	hypothetical prote
92	16	64.0	279	2 T51017	hypothetical prote
93	16	64.0	286	2 G97442	hypothetical prote
94	16	64.0	289	2 H75551	dihydropterolate sy
95	16	64.0	293	2 T31840	hypothetical prote
96	16	64.0	295	2 S48102	xyloglucan endo-1,
97	16	64.0	299	2 F84785	probable transcrip
98	16	64.0	299	2 F83090	hypothetical prote
99	16	64.0	303	2 T29321	hypothetical prote
100	16	64.0	327	2 A46484	apoptosis-mediatin
101	16	64.0	331	2 A71338	conserved hypothet
102	16	64.0	332	2 T26339	hypothetical prote

103	16	64.0	333	2	A96829	probable RING-fing
104	16	64.0	336	2	S76979	hypothetical prote
105	16	64.0	341	2	S63866	platelet activatin
106	16	64.0	341	2	S43352	platelet-activatin
107	16	64.0	348	2	B75445	conserved hypotbet
108	16	64.0	355	2	A34597	rhomboid protein -
109	16	64.0	362	2	T36079	hypothetical prote
110	16	64.0	377	2	T05354	hypothetical prote
111	16	64.0	379	2	S14685	hypothetical prote
112	16	64.0	389	2	T29488	hypothetical prote
113	16	64.0	392	2	I54491	cell surface anti-g
114	16	64.0	416	1	JN0006	neure growth facto
115	16	64.0	418	2	T16713	hypothetical prote
116	16	64.0	429	2	I48361	gene G83-154 prote
117	16	64.0	433	2	P86163	hypothetical prote
118	16	64.0	439	2	G97329	anaerobic dicarbox
119	16	64.0	460	2	T22397	hypothetical prote
120	16	64.0	480	2	T23608	hypothetical prote
121	16	64.0	483	2	T17346	hypothetical prote
122	16	64.0	491	2	T22844	hypothetical prote
123	16	64.0	492	2	S49147	ERf-2 protein - hu
124	16	64.0	496	2	S55665	hypothetical prote
125	16	64.0	521	2	F64522	conserved hypotbet
126	16	64.0	523	2	B71985	hypothetical prote
127	16	64.0	525	2	G71317	hypothetical prote
128	16	64.0	528	2	D70968	hypothetical prote
129	16	64.0	533	2	F72071	conserved hypotbet
130	16	64.0	543	2	C86551	CT384 hypothetical
131	16	64.0	545	2	H83342	periplasmic trehal
132	16	64.0	558	2	A36347	glypican 1 precurs
133	16	64.0	558	2	I56545	glypican precursor
134	16	64.0	561	2	A55235	rhoptyr protein 2
135	16	64.0	579	2	T22703	hypothetical prote
136	16	64.0	583	2	TJ0395	DNA-binding protei
137	16	64.0	586	2	S59944	hydroxymethylgluta
138	16	64.0	603	2	T52507	hypothetical prote
139	16	64.0	604	2	T04120	hydroxymethylgluta
140	16	64.0	618	2	S33044	hypothetical prote
141	16	64.0	625	2	S48941	regulatory protein
142	16	64.0	667	2	A48579	trichozoite surfac
143	16	64.0	687	2	T49226	hypothetical prote
144	16	64.0	692	2	C55926	DNA binding protei
145	16	64.0	700	2	A96690	hypothetical prote
146	16	64.0	713	2	A35502	major surface-labe
147	16	64.0	736	2	JC7518	subtilisin-like se
148	16	64.0	768	2	B41029	integrin beta-8 ch
149	16	64.0	773	1	A36932	iron(II) transport
150	16	64.0	773	2	C91160	ferrous iron trans
151	16	64.0	773	2	B86006	hypothetical prote
152	16	64.0	797	2	A36811	hypothetical prote
153	16	64.0	812	1	PLMS	plasmin (EC 3.4.21
154	16	64.0	829	2	T07406	probable protein K
155	16	64.0	864	2	S61148	SBE2 protein - yea
156	16	64.0	882	2	E96931	hypothetical prote
157	16	64.0	925	2	T02811	DNA excision/repai
158	16	64.0	939	2	C81730	metalloproteinase,
159	16	64.0	942	2	B72015	metalloproteinase,
160	16	64.0	942	2	C86510	insulinase family/
161	16	64.0	953	2	I48078	CHO1 antigen - Chi
162	16	64.0	962	2	JC5571	subtilisin-like pr
163	16	64.0	966	2	A33626	fibrinogen alpha c
164	16	64.0	969	1	A39480	subtilisin-like pr
165	16	64.0	975	2	JC5570	subtilisin-like pr
166	16	64.0	979	2	A89633	protein F13B6.6 [1
167	16	64.0	985	2	S67255	RVF1 protein - yea
168	16	64.0	997	2	A87320	TonB-dependent rec
169	16	64.0	1011	2	S65668	preproteins tranelo
170	16	64.0	1056	2	A53767	mucin MUC5B, trach
171	16	64.0	1058	2	T30178	mitotic checkpoint
172	16	64.0	1102	2	T31004	probable mitotic c
173	16	64.0	1111	2	T01239	hypothetical prote
174	16	64.0	1111	2	T26972	hypothetical prote
175	16	64.0	1151	2	H71347	hypothetical prote

176	16	64.0	1176	2	T49482	hypothetical prote
177	16	64.0	1274	2	T42017	cysteine rich prot
178	16	64.0	1609	2	S25345	probable membrane
179	16	64.0	1729	2	A49282	fuslon protein 1a/
180	16	64.0	1809	2	T49273	hypothetical prote
181	16	64.0	1899	2	T18968	probable serine-ty
182	16	64.0	2105	2	T28669	surface protein 51
183	15	60.0	2233	1	SNMR	metallothionein -
184	15	60.0	25	1	S52359	hypothetical prote
185	15	60.0	25	2	D84641	hypothetical prote
186	15	60.0	43	2	T19161	hypothetical prote
187	15	60.0	44	2	VT9161	hypothetical prote
188	15	60.0	46	1	VT9A1P	viscotoxin 1-PS -
189	15	60.0	46	1	VT9A2	viscotoxin A2 - Eu
190	15	60.0	52	2	JC1197	viscotoxin B - Eur
191	15	60.0	59	1	B44203	metallothionein II
192	15	60.0	60	2	S44350	ferredoxin [4Fe-4S
193	15	60.0	61	1	SMHY1C	ferredoxin [3Fe-4S
194	15	60.0	61	2	I65234	testicular luteal
195	15	60.0	61	2	S00810	metallothionein IC
196	15	60.0	62	2	S15676	chorionic gonadotr
197	15	60.0	62	2	B55995	prostaglandin E2 r
198	15	60.0	67	2	T01958	hypothetical prote
199	15	60.0	77	2	I48725	hypothetical prote
200	15	60.0	78	2	S25619	hypothetical prote
201	15	60.0	78	2	D72764	probable ferredoxi
202	15	60.0	87	2	JT0623	killer toxin HY1 -
203	15	60.0	88	2	JT0514	cysteine-rich 9K p
204	15	60.0	88	2	S12126	cysteine-rich oute
205	15	60.0	88	2	S12125	cysteine-rich oute
206	15	60.0	88	2	E71513	probable 9kDa-cyst
207	15	60.0	88	2	D81671	cysteine rich oute
208	15	60.0	92	1	KVRB38	Ig kappa chain V r
209	15	60.0	93	2	T10379	hypothetical prote
210	15	60.0	95	1	FEDH1	ferredoxin [2Fe-2S
211	15	60.0	95	1	FEDH2	ferredoxin [2Fe-2S
212	15	60.0	97	2	T03100	hypothetical prote
213	15	60.0	101	2	T51856	hypothetical prote
214	15	60.0	101	2	H84599	RING-H2 finger pro
215	15	60.0	103	2	KVRBXP	hypothetical prote
216	15	60.0	104	2	S60405	Ig kappa chain V r
217	15	60.0	104	2	KARB41	Ig kappa chain V r
218	15	60.0	107	1	A72782	Ig kappa chain V r
219	15	60.0	107	2	KVRB12	Ig kappa chain V r
220	15	60.0	108	1	KVRB2K	Ig kappa chain V r
221	15	60.0	108	1	KVRB1	Ig kappa chain V r
222	15	60.0	108	1	KVRB5	Ig kappa chain V r
223	15	60.0	108	2	D85747	hypothetical prote
224	15	60.0	109	1	PVSNNB	Ig kappa chain V r
225	15	60.0	109	1	D71138	hypothetical prote
226	15	60.0	109	2	KVRB29	Ig kappa chain V r
227	15	60.0	110	1	KVRB54	Ig kappa chain V r
228	15	60.0	110	1	S21196	Ig kappa chain V r
229	15	60.0	112	2	A84605	hypothetical prote
230	15	60.0	112	2	MMNVEM	Escort-T medium (ET
231	15	60.0	113	1	PC1274	NS5 protein - hepa
232	15	60.0	113	2	PC1277	NS5 protein - hepa
233	15	60.0	113	2	PC1278	NS5 protein - hepa
234	15	60.0	113	2	PC1276	NS5 protein - hepa
235	15	60.0	113	2	PC1275	NS5 protein - hepa
236	15	60.0	113	2	S57692	probable membrane
237	15	60.0	114	2	KARB2	Ig kappa chain pre
238	15	60.0	117	1	BVECAR	arsenical resistan
239	15	60.0	117	1	G91131	probable truncated
240	15	60.0	117	2	PN0141	lutropin beta chal
241	15	60.0	118	2	PN0139	Iutropin beta chal
242	15	60.0	118	2	B32536	T-cell receptor al
243	15	60.0	118	2	A61465	Iutropin beta chal
244	15	60.0	119	2	D82022	truncated p11n NM
245	15	60.0	120	2	S58424	seminal plasma pro
246	15	60.0	121	2	D69047	ychn protein homol
247	15	60.0	121	2		
248	15	60.0	121	2		



249	15	60.0	122	2	A27635	Ig heavy chain pre	322	15	60.0	189	2	T19559	hypothetical prote
250	15	60.0	122	2	C64965	hypothetical prote	323	15	60.0	190	2	C69345	hypothetical prote
251	15	60.0	122	2	B85825	probable structura	324	15	60.0	191	2	AH3005	conserved hypothet
252	15	60.0	122	2	E90979	probable structura	325	15	60.0	191	2	S69735	hypothetical prote
253	15	60.0	125	2	S53562	probable membrane	326	15	60.0	193	2	E95402	hypothetical prote
254	15	60.0	125	2	B70981	hypothetical prote	327	15	60.0	194	2	UQ1560	hypothetical 20.6k
255	15	60.0	125	2	T01628	hypothetical prote	328	15	60.0	194	2	T41796	hypothetical prote
256	15	60.0	127	2	T44119	hypothetical prote	329	15	60.0	195	2	T07735	nodulin-20a - soyb
257	15	60.0	127	2	B83104	hypothetical prote	330	15	60.0	196	2	PN0478	vancomycin resista
258	15	60.0	127	2	C84302	hypothetical prote	331	15	60.0	197	2	A81877	probable membrane
259	15	60.0	128	2	T18072	hypothetical prote	332	15	60.0	197	2	F72628	hypothetical prote
260	15	60.0	132	2	B24255	chorion class A pr	333	15	60.0	204	2	T28939	hypothetical prote
261	15	60.0	132	2	P00394	genome polypeptin	334	15	60.0	205	2	S37804	hypothetical prote
262	15	60.0	132	2	P00396	genome polypeptin	335	15	60.0	206	2	UC7311	20k protein - soyb
263	15	60.0	134	2	D90229	hypothetical prote	336	15	60.0	206	2	T25139	hypothetical prote
264	15	60.0	135	2	S36197	hypothetical prote	337	15	60.0	208	2	S49196	Kunitz trypsin inh
265	15	60.0	138	2	S00512	luciferin beta chal	338	15	60.0	209	2	A83849	alpha-ribazole-5'-
266	15	60.0	138	2	A82022	truncated pilin NM	339	15	60.0	209	2	S27494	nod8 protein - Bra
267	15	60.0	139	2	S72900	testicular luteinl	340	15	60.0	209	2	E96517	hypothetical prote
268	15	60.0	140	2	H84502	permease r2168b -	341	15	60.0	209	2	T29735	hypothetical prote
269	15	60.0	140	2	UTRBOB	hypothetical prote	342	15	60.0	211	2	C88103	protein M10G11.15
270	15	60.0	141	1	UTRBOB	hypothetical prote	343	15	60.0	212	2	T29906	hypothetical prote
271	15	60.0	141	1	UTRBOB	luciferin beta chal	344	15	60.0	213	2	B25750	nodulin-26b - soyb
272	15	60.0	141	1	UTRBOB	luciferin beta chal	345	15	60.0	213	2	A25750	nodulin-27 - soyb
273	15	60.0	141	1	UTRBOB	luciferin beta chal	346	15	60.0	213	2	A55583	pyrogutamy1-pepti
274	15	60.0	141	1	UTRBOB	luciferin beta chal	347	15	60.0	214	2	D71282	probable deda prot
275	15	60.0	141	2	F82021	luteinizing hormon	348	15	60.0	215	2	C98278	hypothetical prote
276	15	60.0	142	2	JC6305	truncated pilin NM	349	15	60.0	215	2	B57055	melanocortin-4 rec
277	15	60.0	143	2	S52595	neutrin precursor	350	15	60.0	215	2	A25669	nodulin-22 precurs
278	15	60.0	143	2	S52595	probable membrane	351	15	60.0	215	2	T24710	hypothetical prote
279	15	60.0	145	2	T42280	hypothetical prote	352	15	60.0	216	1	F69686	orotate phosphorib
280	15	60.0	145	2	T33037	hypothetical prote	353	15	60.0	216	2	S54430	fimbrial protein h
281	15	60.0	146	2	T01683	immobilization ant	354	15	60.0	217	2	S01358	salivary glue prot
282	15	60.0	150	2	T49563	hypothetical prote	355	15	60.0	220	2	S52005	nodulin-30 (NPV30)
283	15	60.0	150	2	D87652	hypothetical prote	356	15	60.0	220	2	S42879	nodulin-30 - kidne
284	15	60.0	152	2	A32669	vasotocin 1 / neur	357	15	60.0	223	2	A23036	nodulin - soybean
285	15	60.0	154	2	AG0060	probable exported	358	15	60.0	224	2	S07315	nodulin - soybean
286	15	60.0	158	2	H83361	US9 protein PA227	359	15	60.0	225	1	S56809	gene 58 protein -
287	15	60.0	158	2	S35787	vasotocin / neurop	360	15	60.0	225	2	D88492	probable membrane
288	15	60.0	159	2	B29879	vasotocin / neurop	361	15	60.0	225	2	D64598	hypothetical prote
289	15	60.0	160	2	T05385	hypothetical prote	362	15	60.0	225	2	D64598	probable BHLH tran
290	15	60.0	161	1	S43893	H+-exporting ATPas	363	15	60.0	226	2	H96547	hypothetical prote
291	15	60.0	161	2	S14480	arginine-vasococin	364	15	60.0	228	2	B72655	probable transcrip
292	15	60.0	161	2	T13688	desfabiase 2 homo	365	15	60.0	228	2	H71350	hypothetical prote
293	15	60.0	162	2	F97139	hypothetical prote	366	15	60.0	228	2	S26995	cellulose 1,4-beta
294	15	60.0	164	2	A41712	H+-exporting ATPas	367	15	60.0	229	2	S60454	glucose starvation
295	15	60.0	165	2	AH0215	4-hydroxyphenylace	368	15	60.0	230	2	S18989	fama protein - Str
296	15	60.0	165	2	T48806	hypothetical prote	369	15	60.0	231	2	T24669	hypothetical prote
297	15	60.0	169	1	KTHOB	chorigonadotropin	370	15	60.0	234	2	S76342	hypothetical prote
298	15	60.0	170	2	AE0630	4-hydroxyphenylace	371	15	60.0	235	2	F82754	cellulose 1,4-beta
299	15	60.0	170	2	A83135	hypothetical prote	372	15	60.0	236	2	S26993	cellulose 1,4-beta
300	15	60.0	170	2	T06056	hypothetical prote	373	15	60.0	237	1	B42371	regulatory protein
301	15	60.0	172	1	B69138	transcription repr	374	15	60.0	238	1	VOCH	riboflavin-binding
302	15	60.0	173	2	B39141	hypothetical prote	375	15	60.0	238	1	LNRTMA	mannose-binding 1e
303	15	60.0	174	2	C69404	hypothetical prote	376	15	60.0	238	2	A64156	Mg2+-transporting
304	15	60.0	174	2	T49180	hypothetical prote	377	15	60.0	238	2	C83194	conserved hypothet
305	15	60.0	174	2	B85977	hypothetical prote	378	15	60.0	242	2	UC3460	hypothetical prote
306	15	60.0	176	2	E71371	thiamatin-like pro	379	15	60.0	242	2	T16640	ribonuclease (EC 3
307	15	60.0	177	2	S25551	hypothetical prote	380	15	60.0	239	1	JC1373	hypothetical prote
308	15	60.0	179	2	T19557	probable lysine-ke	381	15	60.0	241	2	T30761	hypothetical prote
309	15	60.0	179	2	T19557	hypothetical prote	382	15	60.0	242	2	UC3460	riboflavin-binding
310	15	60.0	180	2	F97324	uncharacterized co	383	15	60.0	242	2	H69525	hypothetical prote
311	15	60.0	180	2	S15327	fimbrial protein M	384	15	60.0	243	2	T31838	hydroxymethylpyrim
312	15	60.0	180	2	C32308	nifQ protein - Rho	385	15	60.0	244	2	T31838	hypothetical prote
313	15	60.0	180	2	P88028	protein P46F5.15 l	386	15	60.0	245	2	T24565	proteolipid protei
314	15	60.0	181	2	T04165	pathogenesis-relat	387	15	60.0	246	2	151325	mannose-binding 1e
315	15	60.0	182	2	T26109	hypothetical prote	388	15	60.0	248	1	LNHUMC	chymotrypsin (EC 3
316	15	60.0	184	2	C72740	hypothetical prote	389	15	60.0	248	2	S49323	uridylylate kinase C
317	15	60.0	185	2	B69028	hypothetical prote	390	15	60.0	248	2	B72045	UMP kinase [import
318	15	60.0	185	2	B26669	probable phosphono	391	15	60.0	248	2	G86577	RING-H2 finger pro
319	15	60.0	186	1	DEPSNL	amine dehydrogenas	392	15	60.0	249	2	T31837	hypothetical prote
320	15	60.0	187	2	T32826	hypothetical prote	393	15	60.0	249	2	T21785	hypothetical prote
321	15	60.0	189	2	S60587	non-structural pro	394	15	60.0	250	1	A49053	CD27 antigen precu

395	15	60.0	250	2	T31836	hypotheoretical prote
396	15	60.0	250	2	T31835	hypotheoretical prote
397	15	60.0	253	2	T49483	hypotheoretical prote
398	15	60.0	255	2	F69962	amino acid ABC tra
399	15	60.0	256	2	B32393	T-cell antigen 4-1
400	15	60.0	257	2	T28946	hypotheoretical prote
401	15	60.0	258	2	G96798	hypotheoretical prote
402	15	60.0	259	2	T52619	TINy-like protein
403	15	60.0	260	1	A46517	CD27 antigen precu
404	15	60.0	262	1	VGIHPC	B1 membrane glycop
405	15	60.0	262	2	S47427	hypotheoretical prote
406	15	60.0	263	2	T30083	hypotheoretical prote
407	15	60.0	266	2	P00393	genome polyprotein
408	15	60.0	266	2	I36857	B27r protein - var
409	15	60.0	268	2	S08229	chlorophyll a/b-bi
410	15	60.0	270	2	I64041	fdhd protein - Hae
411	15	60.0	272	2	T20991	hypotheoretical prote
412	15	60.0	273	2	E75476	hypotheoretical prote
413	15	60.0	275	2	F69229	hypotheoretical prote
414	15	60.0	276	2	AE0860	hypotheoretical prote
415	15	60.0	277	2	S38503	beta-lactamase (EC
416	15	60.0	277	2	S38671	beta-lactamase (EC
417	15	60.0	277	2	I37552	OX40 homolog - hum
418	15	60.0	279	2	B96917	membrane permease,
419	15	60.0	280	2	T02090	xyloglucan endo-1,
420	15	60.0	281	2	E64216	hypotheoretical prote
421	15	60.0	281	2	T03509	probable arsenate
422	15	60.0	285	2	G69204	ribose-phosphate p
423	15	60.0	285	2	T21962	hypotheoretical prote
424	15	60.0	286	2	T02119	probable DnaJ prot
425	15	60.0	289	1	MFHPC	B1 membrane glycop
426	15	60.0	291	2	T28846	hypotheoretical prote
427	15	60.0	292	2	D69176	conserved hypothe
428	15	60.0	294	1	KFM53	lissue factor prec
429	15	60.0	294	2	E84640	CONSTANS-like B-do
430	15	60.0	294	2	T23682	hypotheoretical prote
431	15	60.0	300	2	T05729	hypotheoretical prote
432	15	60.0	303	2	T46715	hypotheoretical prote
433	15	60.0	303	2	T04541	hypotheoretical prote
434	15	60.0	305	2	F82765	conserved hypothe
435	15	60.0	305	2	T25117	hypotheoretical prote
436	15	60.0	307	2	T08125	CONSTANS protein h
437	15	60.0	307	2	D70741	hypotheoretical prote
438	15	60.0	307	2	T18846	hypotheoretical prote
439	15	60.0	312	2	A85778	probable flavoprot
440	15	60.0	312	2	E90529	probable flavoprot
441	15	60.0	312	2	B64928	electron transfer
442	15	60.0	312	2	F69842	3-oxoacyl- acyl-ca
443	15	60.0	312	2	B82339	biotin synthesis p
444	15	60.0	312	2	T31834	hypotheoretical prote
445	15	60.0	312	2	T32379	hypotheoretical prote
446	15	60.0	314	2	I37383	FAS soluble protei
447	15	60.0	314	2	S55179	mtRNA splice defec
448	15	60.0	314	2	G84336	hypotheoretical prote
449	15	60.0	317	2	S29204	melanotropin recep
450	15	60.0	317	2	A64343	hypotheoretical prote
451	15	60.0	318	2	T28941	hypotheoretical prote
452	15	60.0	319	1	B42871	alcohol dehydrogen
453	15	60.0	320	2	S75911	hypotheoretical prote
454	15	60.0	320	2	A53119	cell adhesion glyco
455	15	60.0	321	2	F84597	probable prolina-r
456	15	60.0	321	2	D81018	iron(III) ABC tran
457	15	60.0	321	2	E81962	probable membrane
458	15	60.0	321	2	T24773	hypotheoretical prote
459	15	60.0	321	2	D96035	probable transposa
460	15	60.0	321	2	E96025	probable DNA-(apur
461	15	60.0	323	2	T52563	probable RING zinc
462	15	60.0	324	2	TC3295	probable RING zinc
463	15	60.0	327	2	D86474	hypothetical prote
464	15	60.0	328	2	T19322	salivary glue prot
465	15	60.0	328	2	S01359	heat shock protein
466	15	60.0	328	2	H64554	probable lipid A b
467	15	60.0	328	2	H71954	
468	15	60.0	328	2	G89152	protein C24B5.5 (I
469	15	60.0	330	2	T25169	hypotheoretical prote
470	15	60.0	331	2	T20955	hypotheoretical prote
471	15	60.0	331	2	A54295	interferon alpha/b
472	15	60.0	331	2	S59501	interferon recepto
473	15	60.0	331	2	G96785	protein F10A5.24 (
474	15	60.0	332	2	A57055	melanocortin recep
475	15	60.0	333	2	A47536	gene WNT3 protein
476	15	60.0	335	2	S52890	hypotheoretical prote
477	15	60.0	335	2	A40036	apoptosis-mediati
478	15	60.0	335	2	UB0115	zinc-finger protei
479	15	60.0	336	2	A46478	major surface anti
480	15	60.0	336	2	A30527	interferon gamma r
481	15	60.0	337	2	I38500	hypotheoretical prote
482	15	60.0	339	2	H71265	GTP-binding regula
483	15	60.0	340	1	RGHDB3	G-protein beta-sub
484	15	60.0	340	2	I53871	DNA-binding protei
485	15	60.0	341	2	S72445	vancomycin resist
486	15	60.0	343	2	JN0249	38.7K hypotheical
487	15	60.0	343	2	H86385	hypotheical prote
488	15	60.0	344	2	F70922	NADH2 dehydrogena
489	15	60.0	345	1	QXXL2M	surface antigen 51
490	15	60.0	345	2	A05279	hypotheical prote
491	15	60.0	346	2	T27896	probable Fe(II) tr
492	15	60.0	346	2	T06385	iron(III) diclitrat
493	15	60.0	348	2	S74439	interferon recepto
494	15	60.0	349	2	JC6311	hypotheical prote
495	15	60.0	350	2	S21400	wingless homolog X
496	15	60.0	352	2	A48828	Wnt-3A protein - m
497	15	60.0	352	2	A39532	conserved hypothe
498	15	60.0	352	2	E95381	hypotheical prote
499	15	60.0	352	2	G71914	P2455.9 [imported]
500	15	60.0	353	2	C96711	IL1 protein - rheu
501	15	60.0	354	1	PIWLR1	Wnt-3 protein - mo
502	15	60.0	355	2	A35503	CONSTANS-like 1 -
503	15	60.0	355	2	T51414	hypotheical prote
504	15	60.0	355	2	T29932	hypotheical prote
505	15	60.0	356	2	T20656	inorganic diphosph
506	15	60.0	360	2	S54173	nodulin - soybean
507	15	60.0	360	2	S09552	ABC-transporter At
508	15	60.0	361	2	T34595	hypotheical prote
509	15	60.0	361	2	T37938	cyst wall protein
510	15	60.0	362	2	S61924	hypotheical 38.5
511	15	60.0	363	2	E65104	ABC transporter ph
512	15	60.0	364	2	G82595	hypotheical prote
513	15	60.0	364	2	T30723	hypotheical prote
514	15	60.0	364	2	B85430	NS5 protein - hepa
515	15	60.0	365	2	JU0879	probable alcohol d
516	15	60.0	366	1	A30196	DNA-directed DNA p
517	15	60.0	366	1	A64107	phosphate transpor
518	15	60.0	366	2	T01169	hypotheical prote
519	15	60.0	366	2	F72062	CT449 hypotheical
520	15	60.0	366	2	A86561	prephenate dehydra
521	15	60.0	368	2	S59565	probable adenosylm
522	15	60.0	369	2	T10750	com3 protein - My
523	15	60.0	369	2	S73716	protein phosphatas
524	15	60.0	370	2	S54297	B2 protein - human
525	15	60.0	372	1	W2WL31	alcohol dehydrogen
526	15	60.0	372	2	AE3184	hypotheical prote
527	15	60.0	372	2	T23680	hypotheical prote
528	15	60.0	374	2	T22149	hypotheical trans
529	15	60.0	376	2	S70841	beta-lactamase (EC
530	15	60.0	377	1	QKEC	beta-lactamase (EC
531	15	60.0	377	2	T12042	cysteine proteinas
532	15	60.0	377	2	C86111	beta-lactamase, pe
533	15	60.0	377	2	C91270	protein R09F10.7 (
534	15	60.0	380	2	B89588	hypotheical prote
535	15	60.0	380	2	H70590	lipase homolog T29
536	15	60.0	381	2	T06696	hypotheical prote
537	15	60.0	382	2	T25280	carbazole dioxygen
538	15	60.0	384	2	T46578	hypotheical prote
539	15	60.0	385	2	T31493	
540	15	60.0	389	2	T23167	

541	15	60.0	390	2	D97305	NADH-dependent but	614	15	60.0	485	2	T13479	hypothetical prote
542	15	60.0	390	2	B47013	butanol dehydrogen	615	15	60.0	487	2	T13875	hypothetical prote
543	15	60.0	392	2	C82628	hypothetical prote	616	15	60.0	490	2	T23003	hypothetical prote
544	15	60.0	392	2	T16776	hypothetical prote	617	15	60.0	493	2	S36931	chitinase (EC 3.2-
545	15	60.0	393	2	T16046	hypothetical prote	618	15	60.0	498	2	A48203	interleukin-14 pre
546	15	60.0	393	2	S48288	probable phosphopr	619	15	60.0	500	2	A71306	hypothetical prote
547	15	60.0	394	2	H72505	hypothetical prote	620	15	60.0	501	2	C84771	probable glucosyl
548	15	60.0	394	2	F95973	probable sugar upt	621	15	60.0	501	2	S43536	T-cell surface gly
549	15	60.0	396	1	TRXXB4	alpha-lytic proteol	622	15	60.0	501	2	JH0447	alpha-1A-adrenorgi
550	15	60.0	396	1	DCBYDM	adenosylmethionine	623	15	60.0	501	2	E96758	hypothetical prote
551	15	60.0	396	1	E64987	bicyclomyacin resis	624	15	60.0	504	1	B48939	cellulose 1,4-beta
552	15	60.0	396	2	B91013	bicyclomyacin resis	625	15	60.0	507	2	S18450	variant surface gl
553	15	60.0	396	2	D85857	bicyclomyacin resis	626	15	60.0	509	1	S14629	aldhyde dehydroge
554	15	60.0	398	2	F40511	hypothetical prote	627	15	60.0	510	2	S41943	cellulose 1,4-beta
555	15	60.0	398	2	S17428	interleukin-1 rece	628	15	60.0	510	2	A70580	probable UDP-N-ace
556	15	60.0	400	2	JC4591	alpha-1,3 fucosylt	629	15	60.0	511	2	S44716	cellulose 1,4-beta
557	15	60.0	400	2	T32705	hypothetical prote	630	15	60.0	512	2	A87270	hypothetical prote
558	15	60.0	405	2	B36340	alpha(1,3)-fucosyl	631	15	60.0	513	2	T37612	cytosol aminopepti
559	15	60.0	406	2	T01931	adenyl cyclase -	632	15	60.0	515	2	A40491	alpha-1B-adrenorgi
560	15	60.0	407	2	JR0113	zinc-finger protei	633	15	60.0	515	2	JC1525	interferon alpha-b
561	15	60.0	410	2	JC7757	function-unknown p	634	15	60.0	516	2	S33164	cellulose 1,4-beta
562	15	60.0	411	2	T29475	hypothetical prote	635	15	60.0	516	2	J50083	cellulose 1,4-beta
563	15	60.0	414	2	H84596	hypothetical prote	636	15	60.0	516	2	A45121	alpha-1B adrenorgi
564	15	60.0	416	2	I45890	ferrochelatase (EC	637	15	60.0	517	2	B86299	hypothetical prote
565	15	60.0	417	2	B86919	probable seryl-trN	638	15	60.0	518	2	RD1NTS	dihydrofolate redu
566	15	60.0	417	2	E30341	alpha-1-adrenorgic	639	15	60.0	520	1	RD1NTS	dihydrofolate redu
567	15	60.0	417	2	T20199	hypothetical prote	640	15	60.0	520	2	T20007	hypothetical prote
568	15	60.0	418	2	E71716	proline/betaine tr	641	15	60.0	520	2	S14600	E2 glycoprotein pr
569	15	60.0	423	2	A36794	hypothetical prote	642	15	60.0	521	2	E82508	metyl-accepting c
570	15	60.0	423	2	T19581	hypothetical prote	643	15	60.0	522	2	B81679	conserved hypotet
571	15	60.0	425	2	D88115	protein F53C3.11 f	644	15	60.0	526	2	S62194	hypothetical prote
572	15	60.0	425	2	C89753	protein F1IC7.3 f1	645	15	60.0	529	2	A45392	RNA-directed RNA p
573	15	60.0	427	2	G88492	protein T07E3.4 f1	646	15	60.0	533	2	B56110	tyrosine phosphopr
574	15	60.0	429	1	T03507	probable carboxype	647	15	60.0	533	2	A83632	D-ribulokinase (EC
575	15	60.0	431	2	T14414	S-locus-specific g	648	15	60.0	538	2	D71522	hypothetical prote
576	15	60.0	432	2	A83616	dicarboxylic acid	649	15	60.0	539	2	S41942	cellulose 1,4-beta
577	15	60.0	432	2	T34154	hypothetical prote	650	15	60.0	540	2	B64829	membrane protein p
578	15	60.0	433	2	T14329	dermal glycoprotei	651	15	60.0	540	2	H90751	probable transport
579	15	60.0	433	2	A57596	alpha-1,3-fucosylt	652	15	60.0	540	2	P85615	hypothetical prote
580	15	60.0	434	2	G86163	hypothetical prote	653	15	60.0	541	2	D84681	hypothetical prote
581	15	60.0	436	1	S49458	diphosphate-fructo	654	15	60.0	540	1	VG1HD6	E2 glycoprotein pr
582	15	60.0	436	1	D69270	ferredoxin-nitrite	655	15	60.0	550	1	S47034	cell fusion protei
583	15	60.0	438	1	ERADC2	early E1B 49K prot	656	15	60.0	552	2	I46329	cell fusion glycop
584	15	60.0	439	2	T15748	hypothetical prote	657	15	60.0	553	1	A63329	cell fusion glycop
585	15	60.0	445	2	A48881	serotonin receptor	658	15	60.0	553	1	VGNZGB	cell fusion glycop
586	15	60.0	445	2	T21744	serotonin receptor	659	15	60.0	553	1	A36830	cell fusion glycop
587	15	60.0	448	2	S36402	serotonin receptor	660	15	60.0	553	1	A46329	cell fusion glycop
588	15	60.0	448	2	A47519	serotonin receptor	661	15	60.0	553	1	B36830	cell fusion glycop
589	15	60.0	449	1	A42800	beta-alanine-pyruv	662	15	60.0	553	1	B46329	cell fusion glycop
590	15	60.0	450	2	E96695	hypothetical prote	663	15	60.0	553	1	D46329	cell fusion glycop
591	15	60.0	450	2	B36810	hypothetical prote	664	15	60.0	553	1	E46329	cell fusion glycop
592	15	60.0	451	2	T23731	hypothetical prote	665	15	60.0	553	1	G46329	cell fusion glycop
593	15	60.0	454	2	T26654	hypothetical prote	666	15	60.0	553	1	VGNZND	cell fusion glycop
594	15	60.0	456	2	T16796	hypothetical prote	667	15	60.0	553	1	VGNZNV	cell fusion glycop
595	15	60.0	458	2	B83341	L-serine dehydrata	668	15	60.0	553	1	VGNZTE	cell fusion glycop
596	15	60.0	458	2	T16123	hypothetical prote	669	15	60.0	553	1	VGNZU1	cell fusion glycop
597	15	60.0	460	2	C82399	conserved hypotet	670	15	60.0	553	2	S06345	cell fusion glycop
598	15	60.0	461	1	A35356	tumor necrosis fac	671	15	60.0	553	2	S38784	gene F protein - N
599	15	60.0	463	1	T49373	hypothetical prote	672	15	60.0	553	2	S40163	cell fusion protei
600	15	60.0	469	1	NM1V27	exo-alpha-sialidas	673	15	60.0	553	2	S38786	gene F protein - N
601	15	60.0	469	1	NM1V22	exo-alpha-sialidas	674	15	60.0	553	2	S23621	gene F protein - N
602	15	60.0	469	1	NM1V2	exo-alpha-sialidas	675	15	60.0	553	2	S23620	F protein - Newcas
603	15	60.0	469	1	J01644	exo-alpha-sialidas	676	15	60.0	553	2	S38785	gene F protein - N
604	15	60.0	470	2	S78440	phosphoglucumutase	677	15	60.0	553	2	T15094	hypothetical prote
605	15	60.0	471	2	S36655	UDP Thamnose-antio	678	15	60.0	553	2	S23622	gene F protein - N
606	15	60.0	473	2	S60290	anthocyanin rhanno	679	15	60.0	554	2	B82934	hypothetical prote
607	15	60.0	473	2	AG0612	probable transport	680	15	60.0	554	2	A31300	perforin precursor
608	15	60.0	474	1	VYHUD	vitamin D-binding	681	15	60.0	554	2	A45818	cytolysin precursor
609	15	60.0	476	2	S09152	translation elonga	682	15	60.0	555	2	A37181	perforin 1 precurs
610	15	60.0	477	2	S71323	alpha-1A adrenorgi	683	15	60.0	555	2	S63137	hypothetical prote
611	15	60.0	481	2	A89102	protein P25E5.4 f1	684	15	60.0	557	2	A48434	variant-specific s
612	15	60.0	483	2	E84454	hypothetical prote	685	15	60.0	560	2	A38731	alpha-1A adrenorgi
613	15	60.0	483	2	T20895	hypothetical prote	686	15	60.0	563	2	C86397	protein TM9.10 f1

687	15	60.0	567	2	D83400	conserved hypothe	760	15	60.0	737	2	AH1958	hypotheical prote
688	15	60.0	568	2	E96648	hypotheical prote	761	15	60.0	739	2	T12964	subtilisin homolog
689	15	60.0	571	2	S69310	protein kinase cak	762	15	60.0	740	2	T03847	fas-binding protei
690	15	60.0	572	2	I39369	alpha-1A-adrenergi	763	15	60.0	753	2	B36268	platelet glycoprot
691	15	60.0	573	2	T29880	hypotheical prote	764	15	60.0	753	2	T28787	hypotheical prote
692	15	60.0	572	2	T25397	hypotheical prote	765	15	60.0	761	2	UC5759	brain-specific ser
693	15	60.0	580	2	S72211	N-acetyl-beta-D-gl	766	15	60.0	766	2	S61420	inorganic diphosph
694	15	60.0	581	2	S17150	potassium channel	767	15	60.0	770	1	A38230	LDL receptor-relat
695	15	60.0	590	2	D84523	probable auxin-reg	768	15	60.0	770	2	T00204	LDL receptor-relat
696	15	60.0	590	2	A45283	interferon alpha/b	769	15	60.0	770	2	T00203	protein PgC16.25 l
697	15	60.0	590	2	A48461	ovarian abundant m	770	15	60.0	772	2	D96504	platelet glycoprot
698	15	60.0	597	2	JU0107	serum albumin 66k p	771	15	60.0	778	2	A60798	ribonucleoside-dip
699	15	60.0	600	2	A47391	serum albumin prec	772	15	60.0	780	2	T03156	integrin beta1 - c
700	15	60.0	603	2	S28941	coagulation factor	773	15	60.0	781	2	S43534	hypothetical 82k p
701	15	60.0	605	1	ABPCS	serum albumin prec	774	15	60.0	784	2	JU0317	hypothetical 82k p
702	15	60.0	605	2	T34040	hypothetical prote	775	15	60.0	788	2	A26547	platelet glycoprot
703	15	60.0	608	1	ABONS1	serum albumin 1 pr	776	15	60.0	788	2	I77349	platelet glycoprot
704	15	60.0	608	1	ABONS2	serum albumin 2 pr	777	15	60.0	788	2	I51530	integrin beta-3 su
705	15	60.0	608	1	ABRS	serum albumin prec	778	15	60.0	792	2	T21276	hypothetical prote
706	15	60.0	608	2	S57632	serum albumin prec	779	15	60.0	792	2	H82406	probable lipase VC
707	15	60.0	608	2	T19935	hypothetical prote	780	15	60.0	796	2	UC7355	peroxisome prolife
708	15	60.0	609	1	ABHUS	hypothetical prote	781	15	60.0	796	2	T16424	hypothetical prote
709	15	60.0	610	2	S35049	serum albumin prec	782	15	60.0	803	2	A47723	F-opsidin precurs
710	15	60.0	611	2	D96607	mucin JBR57 - huma	783	15	60.0	811	2	T25166	hypothetical prote
711	15	60.0	612	2	T05750	hypothetical prote	784	15	60.0	813	2	T21192	DNA polymerase III
712	15	60.0	612	2	S55084	auxin-regulated pr	785	15	60.0	815	2	T05555	hypothetical prote
713	15	60.0	613	2	A39402	probable membrane	786	15	60.0	818	2	T32154	hypothetical prote
714	15	60.0	613	2	T00853	potassium channel	787	15	60.0	819	2	C84615	hypothetical prote
715	15	60.0	614	2	E84006	hypothetical prote	788	15	60.0	822	2	T01095	hypothetical prote
716	15	60.0	618	2	T00476	formate dehydrogen	789	15	60.0	825	2	T29634	hypothetical prote
717	15	60.0	619	1	JH0776	hydrogenase (EC 1.	790	15	60.0	827	1	S28273	diacylglycerol kin
718	15	60.0	619	2	T08613	hypothetical prote	791	15	60.0	828	2	S52393	beta-galactosidase
719	15	60.0	620	2	T27008	hypothetical prote	792	15	60.0	833	2	T10695	transcription fact
720	15	60.0	621	2	S72493	lactase (EC 1.10.3	793	15	60.0	834	2	T28250	ORF MSV089 probab
721	15	60.0	621	2	UC7278	adaptor protein co	794	15	60.0	836	2	E96561	hypothetical prote
722	15	60.0	621	2	T22904	hypothetical prote	795	15	60.0	840	2	H86429	hypothetical prote
723	15	60.0	624	2	S22703	voltage-gated pote	796	15	60.0	842	2	T05400	hypothetical prote
724	15	60.0	624	2	G87619	sensor histidine k	797	15	60.0	851	2	S12159	env protein - huma
725	15	60.0	626	2	T04895	vacuolar sorting r	798	15	60.0	852	1	GNNYHA	env polypeptotein
726	15	60.0	627	2	H82573	exonuclease ABC s	799	15	60.0	852	1	VCLJGG	env polypeptotein
727	15	60.0	627	2	B48442	membrane transport	800	15	60.0	852	2	T01364	homeodomain transc
728	15	60.0	627	4	A40201	artifact-warning s	801	15	60.0	858	1	VCLJG2	env polypeptotein
729	15	60.0	629	2	S60385	probable membrane	802	15	60.0	859	1	VCLJST	env polypeptotein
730	15	60.0	630	2	T43460	hypothetical prote	803	15	60.0	859	1	VCLJCT	env polypeptotein
731	15	60.0	630	2	A54659	DNA repair protei	804	15	60.0	859	1	VCLJ22	env polypeptotein
732	15	60.0	631	2	T31782	hypothetical prote	805	15	60.0	859	1	VCLJEB	env polypeptotein
733	15	60.0	631	2	T29926	hypothetical prote	806	15	60.0	859	1	VCLJRE	env polypeptotein
734	15	60.0	632	2	S42731	collagen alpha 1 c	807	15	60.0	859	1	VCLJEB	env polypeptotein
735	15	60.0	632	2	A71259	probable dicarboxy	808	15	60.0	859	1	VCLJEW	env polypeptotein
736	15	60.0	633	2	A36353	DNA repair protei	809	15	60.0	859	1	VCLJEW	env polypeptotein
737	15	60.0	633	2	T64143	hypothetical prote	810	15	60.0	859	2	S24571	env protein - huma
738	15	60.0	640	2	S69546	phosphoenolpyruvat	811	15	60.0	859	2	C96504	protein PgC16.23 l
739	15	60.0	654	2	H86373	protein T23823.16	812	15	60.0	860	2	T37768	probable vacuolar
740	15	60.0	654	2	B75587	probable N-glycosyl	813	15	60.0	860	2	A48825	Notch homolog Motc
741	15	60.0	663	2	B65025	hypothetical prote	814	15	60.0	861	2	S64821	probable membrane
742	15	60.0	663	2	A91048	hypothetical prote	815	15	60.0	862	2	D88465	protein B0244.7 l1
743	15	60.0	664	2	T06598	probable 2-compone	816	15	60.0	863	2	D88465	hypothetical prote
744	15	60.0	665	1	A42792	snak-type molecula	817	15	60.0	864	2	B86449	hypothetical prote
745	15	60.0	670	2	T32221	succinate dehydrog	818	15	60.0	869	2	S53098	envelope polypept
746	15	60.0	678	2	H96552	hypothetical prote	819	15	60.0	870	2	A96637	hypothetical prote
747	15	60.0	679	2	E85892	hypothetical prote	820	15	60.0	884	2	T18649	hypothetical prote
748	15	60.0	692	2	S37159	NADPH-ferrihemopro	821	15	60.0	890	2	F84548	hypothetical prote
749	15	60.0	695	2	S49165	transferrin precur	822	15	60.0	892	2	A41697	nitrate assimilact
750	15	60.0	699	2	A43734	probable protein k	823	15	60.0	909	2	H87729	protein Y23H5A.7 l
751	15	60.0	703	2	B82148	ATP-dependent heli	824	15	60.0	915	2	T03589	probable aspartate
752	15	60.0	704	2	B86855	hypothetical prote	825	15	60.0	915	2	T16623	hypothetical prote
753	15	60.0	704	2	T38117	probable protein k	826	15	60.0	923	2	E83574	lipoprotein recept
754	15	60.0	708	2	T19417	hypothetical prote	827	15	60.0	923	2	T37475	hypothetical prote
755	15	60.0	711	2	T30107	hypothetical prote	828	15	60.0	929	2	T32492	hypothetical prote
756	15	60.0	712	1	VCLJ84	env polypeptotein	829	15	60.0	931	2	T33744	hypothetical prote
757	15	60.0	720	2	A36526	choline O-acetyltr	830	15	60.0	942	2	C96574	hypothetical prote
758	15	60.0	728	2	T20561	hypothetical prote	831	15	60.0	946	2	T00024	ent-kaurene syntha
759	15	60.0	736	2	T12963	subtilisin homolog	832	15	60.0	959	2	E85276	hypothetical prote



979 14 56.0 47 2 S22828  
 980 14 56.0 47 2 I48948  
 981 14 56.0 49 2 S00228  
 982 14 56.0 49 2 S02007  
 983 14 56.0 50 1 HSPG  
 984 14 56.0 50 1 HSPG  
 985 14 56.0 50 2 S22582  
 986 14 56.0 50 2 S21672  
 987 14 56.0 50 2 T38209  
 988 14 56.0 51 1 HSBOS  
 989 14 56.0 51 1 HSBOS1  
 990 14 56.0 52 2 S03997  
 991 14 56.0 52 2 S65712  
 992 14 56.0 53 2 S24596  
 993 14 56.0 53 2 S75905  
 994 14 56.0 55 1 B7AG58  
 995 14 56.0 56 1 NIBSSA  
 996 14 56.0 56 2 T12783  
 997 14 56.0 57 2 S12957  
 998 14 56.0 58 1 SMK15  
 999 14 56.0 58 2 G97810  
 1000 14 56.0 59 1 FEDVIV

## ALIGNMENTS

RESULT 1  
 T28945  
 hypothetical protein F07C4.2 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T28945  
 R/Miller, N.; Steilyes, L.  
 submitted to the EMBL Data Library, January 1997  
 A/Description: The sequence of C. elegans cosmid F07C4.  
 A/Reference number: Z20546  
 A/Accession: T28945  
 A/Status: preliminary; translated from GB/EMBL/DBD  
 A/Molecule type: DNA  
 A/Residues: 1-155 <MIL>  
 A/Cross-references: UNIPROT:P91214; UNIPARC:UPI000007DB2D; EMBL:U80023; PIDN:AAC48012.1  
 A/Experimental source: strain Bristol N2; clone F07C4  
 C/Genetics:  
 A/Gene: CSEP:F07C4.2  
 A/Map position: 5  
 A/Intons: 116/1  
 C/Superfamily: tetranectin; C-type lectin homology

Query Match 72.0%; Score 18; DB 2; Length 155;  
 Best Local Similarity 22.2%; Pred. No. 17;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 92 CTTASSSC 100

RESULT 2  
 T45962  
 hypothetical protein F7J8.200 - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
 C/Accession: T45962  
 R/Bevan, M.; Zimmermann, W.; Gueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Lem  
 submitted to the Protein Sequence Database, January 2000  
 A/Reference number: Z23018  
 A/Accession: T45962  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-470 <BRV>  
 A/Cross-references: UNIPROT:Q9LFB4; UNIPARC:UPI00000A7D3B; EMBL:AL137189  
 A/Experimental source: cultivar Columbia; BAC clone F7J8

C/Genetics:  
 A/Map position: 5  
 A/Intons: 116/2; 138/3; 162/2; 201/3; 226/2; 250/1; 275/1; 288/2; 307/2; 379/3  
 A/Note: F7J8.200

Query Match 72.0%; Score 18; DB 2; Length 470;  
 Best Local Similarity 22.2%; Pred. No. 23;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 23 CSSSSATSC 31

RESULT 3  
 A34598  
 eclystone-induced protein E75A - fruit fly (Drosophila melanogaster)  
 C/Species: Drosophila melanogaster  
 C/Date: 29-Jun-1990 #sequence\_revision 29-Jun-1990 #text\_change 31-Dec-2004  
 C/Accession: A34598  
 R/Segraves, W.A.; Hogness, D.S.  
 Genes Dev. 4, 204-219, 1990  
 A/Title: The E75 eclystone-inducible gene responsible for the 75B early puff in Drosophila  
 A/Reference number: A34598; M0ID:90249727; PMID:2110921  
 A/Accession: A34598  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-1237 <SEG>  
 A/Cross-references: UNIPROT:P17671; UNIPARC:UPI0000129B0; GB:X51548; NID:g7910; PIDN:CA  
 C/Genetics:  
 A/Gene: FlyBase:Elp75B  
 A/Cross-references: FlyBase:FBgn000568  
 C/Keywords: alternative splicing; DNA binding; nucleus; transcription regulation; zinc f  
 P;243-520/Domain: erba transforming protein homology <ERBA>

Query Match 72.0%; Score 18; DB 2; Length 1237;  
 Best Local Similarity 22.2%; Pred. No. 30;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 1070 CSSSSASSC 1078

RESULT 4  
 B34598  
 eclystone-induced protein E75B - fruit fly (Drosophila melanogaster)  
 C/Species: Drosophila melanogaster  
 C/Date: 29-Jun-1990 #sequence\_revision 29-Jun-1990 #text\_change 31-Dec-2004  
 C/Accession: B34598  
 R/Segraves, W.A.; Hogness, D.S.  
 Genes Dev. 4, 204-219, 1990  
 A/Title: The E75 eclystone-inducible gene responsible for the 75B early puff in Drosophila  
 A/Reference number: A34598; M0ID:90249727; PMID:2110921  
 A/Accession: B34598  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-1394 <SEG>  
 A/Cross-references: UNIPROT:P17672; UNIPARC:UPI0000129B0; GB:X51549; NID:g7912; PIDN:CA  
 C/Genetics:  
 A/Gene: FlyBase:Elp75B  
 A/Cross-references: FlyBase:FBgn000568  
 C/Keywords: alternative splicing; DNA binding; nucleus; transcription regulation; zinc f  
 F;401-677/Domain: erba transforming protein homology <ERBA>

Query Match 72.0%; Score 18; DB 2; Length 1394;  
 Best Local Similarity 22.2%; Pred. No. 31;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 1227 CSSSSASSC 1235

## RESULT 5

S05979  
steroid hormone receptor homolog - fruit fly (Drosophila melanogaster)  
N:Alternate names: puff 75B protein  
C:Species: Drosophila melanogaster  
C>Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 05-Oct-2004  
C:Accession: S05979  
R:Feigl, G.; Gram, M.; Pongs, O.  
Nucleic Acids Res. 17, 7167-7178, 1989  
A:Title: A member of the steroid hormone receptor gene family is expressed in the 20-OH-  
A:Reference number: S05979; M01D:90016778; PMID:2508058  
A:Accession: S05979  
A:Molecule type: mRNA  
A:Residues: 1-1443 <FEI>  
A:Cross-references: UNIPROT: P13055; UNIPARC: UP10000129B02; EMBL: X15586; NID: g7516; PIDN:  
C:Genetics:  
A:Gene: FlyBase: FlyBase: FBgn0000568  
A:Cross-references: FlyBase: FBgn0000568  
A:Map position: 3 75B  
C:Keywords: alternative splicing; DNA binding; steroid hormone receptor; transcription  
F:452-727/Domain: erba transforming protein homology <ERBA>

Query Match 72.0%; Score 18; DB 2; Length 1443;  
Best Local Similarity 22.2%; Pred. No. 32;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 9  
DB 1276 CSSSASASC 1284

## RESULT 6

TJ3355  
hypothetical protein F16G10.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: TJ3355  
R:Gutting, S.; Scheet, P.  
Submitted to the EMBL Data Library, July 1998  
A:Description: The sequence of C. elegans cosmid F16G10.  
A:Reference number: 221329  
A:Accession: TJ3355  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-199 <GAT>  
A:Cross-references: UNIPROT: O76588; UNIPARC: UP1000007891D; EMBL: AF077537; PIDN: AAC26272.  
A:Experimental source: strain Bristol N2; clone F16G10  
C:Genetics:  
A:Gene: CESP: F16G10.4  
A:Map position: 2  
A:Intons: 56/3; 69/1; 89/1; 133/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein F16G10.4

Query Match 68.0%; Score 17; DB 2; Length 199;  
Best Local Similarity 22.2%; Pred. No. 51;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 9  
DB 72 CEATSTATC 80

## RESULT 7

S63342  
hypothetical protein YNR014w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein N2060  
C:Species: Saccharomyces cerevisiae  
C>Date: 27-Apr-1996 #sequence\_revision 03-May-1996 #text\_change 09-Jul-2004  
C:Accession: S63342; S63344; S63340  
R:Maurer, C.T.C.; Urbanus, J.H.M.; Plant, R.U.  
Submitted to the Protein Sequence Database, April 1996  
A:Reference number: S63266

A:Accession: S63342  
A:Molecule type: DNA  
A:Residues: 1-212 <MAU>  
A:Cross-references: UNIPROT: P53719; UNIPARC: UP1000013BAB3; EMBL: Z71629; NID: g1302493; PI  
A:Experimental source: strain S288C  
R:Duetsch, A.; Floeth, M.; Fritz, C.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D.  
Submitted to the Protein Sequence Database, April 1996  
A:Reference number: S62944  
A:Accession: S63344  
A:Molecule type: DNA  
A:Residues: 1-212 <DUB>  
A:Cross-references: UNIPARC: UP1000013BAB3; EMBL: Z71629; NID: g1302493; PID: e239786; PID: g

A:Experimental source: strain S288C  
R:Aert, R.; Verhaesselt, P.; Voet, M.; Volkert, G.  
Submitted to the Protein Sequence Database, April 1996  
A:Reference number: S62910  
A:Accession: S63340  
A:Molecule type: DNA  
A:Residues: 1-26 <AER>  
A:Cross-references: UNIPARC: UP10000179C25; EMBL: Z71629; MIPS: YNR014w  
A:Experimental source: strain S288C  
C:Genetics:  
A:Cross-references: SGD: S0005297  
A:Map position: 14R

Query Match 68.0%; Score 17; DB 2; Length 212;  
Best Local Similarity 22.2%; Pred. No. 52;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 9  
DB 40 CSSATSAC 48

## RESULT 8

C70812  
probable lpgQ protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: C70812  
R:Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentile, S.; Hamlin, N.; Holroyd, S.  
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; M01D: 98295987; PMID: 9634230  
A:Accession: C70812  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-214 <COL>  
A:Cross-references: UNIPROT: O53846; UNIPARC: UP10000165256; GB: AL022004; GB: AL123456; NIT  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: lpgQ

Query Match 68.0%; Score 17; DB 2; Length 214;  
Best Local Similarity 22.2%; Pred. No. 52;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 9  
DB 15 CAIAATTAC 23

## RESULT 9

H72532  
hypothetical protein APE2236 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: H72532  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin, K.; Takai  
awa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;

DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum  
 A:Reference number: A72450, MUID:99310339, PMID:10382966  
 A:Accession: H72532  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-244 <RAM>  
 A:Cross-references: UNIPROT:Q9Y9Q2; UNIPARC:UPI000005E27C; DDBJ:AP000063; NID:95105654;  
 A:Experimental source: strain KI  
 A:Genetics:  
 A:Gene: APE2236  
 C:Superfamily: Aeropyrum pernix hypothetical protein APE2236

Query Match 68.0%; Score 17; DB 2; Length 244;  
 Best Local Similarity 22.2%; Pred. No. 54;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 DB 107 CSSSSAALC 115

RESULT 10  
 T40415  
 hypothetical protein SPBC4.01 - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C:Accession: T40415  
 R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Harris, D.  
 submitted to the EMBL Data Library, October 1999  
 A:Reference number: Z21927  
 A:Accession: T40415  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-248 <MCD>  
 A:Cross-references: UNIPROT:Q9US9; UNIPARC:UPI000006AF49; EMBL:AL121863; PIDN:CA858400;  
 A:Experimental source: strain 972h-; coemid c4  
 C:Genetics:  
 A:Gene: SPDB:SPBC4.01  
 A:Map position: 2  
 A:introns: 230/2

Query Match 68.0%; Score 17; DB 2; Length 248;  
 Best Local Similarity 22.2%; Pred. No. 54;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 DB 146 CLTTSTAAC 154

RESULT 11  
 H71353  
 conserved hypothetical protein TP0182 - syphilis spirochete  
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
 C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
 C:Accession: H71353  
 R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Ueberback, T.; MDO  
 they, L.; Weisman, J.; Smith, H.O.; Venter, J.C.  
 Science 281, 375-388, 1998  
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
 A:Reference number: A71250, MUID:98332770, PMID:9665876  
 A:Accession: H71353  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-259 <COL>  
 A:Cross-references: UNIPROT:Q83212; UNIPARC:UPI000000CA51; GB:AE001202; GB:AE000520; NID  
 A:Experimental source: strain Nichols  
 C:Genetics:  
 A:Gene: TP0182  
 C:Superfamily: RNA-binding protein, YrDC type

Query Match 68.0%; Score 17; DB 2; Length 259;  
 Best Local Similarity 22.2%; Pred. No. 55;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 DB 31 CASASLISC 39

RESULT 12  
 AC0661  
 probable lipoprotein STY1394 [imported] - Salmonella enterica subsp. enterica serovar Ty  
 C:Species: Salmonella enterica subsp. enterica serovar Typh  
 A:Note: this species has also been called Salmonella typhi  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C:Accession: AC0661  
 R:Fairhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
 th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
 S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
 A:Reference number: AB0502; MUID:21534947; PMID:11677608  
 A:Accession: AC0661  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-281 <PAR>  
 A:Cross-references: UNIPARC:UPI0000059F0E; GB:AL513382; PIDN:CAD01660.1; PID:gl6502512;  
 C:Genetics:  
 A:Gene: STY1394

Query Match 68.0%; Score 17; DB 2; Length 281;  
 Best Local Similarity 22.2%; Pred. No. 56;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 DB 208 CAATAAEC 216

RESULT 13  
 T48975  
 xyloglucan endo-transglycosylase - Arabidopsis thaliana  
 N:Alternate names: protein FlAD17.60  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
 C:Accession: T48975  
 R:Jordan, N.; Bangert, S.; Wiedemann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Rudd, S.;  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z25008  
 A:Accession: T48975  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-293 <TOR>  
 A:Cross-references: UNIPROT:P93046; UNIPARC:UPI00000487CE; EMBL:AJ353992; GSPDB:GN00061,  
 A:Experimental source: cultivar Columbia; BAC clone FlAD17  
 C:Genetics:  
 A:Gene: ATSP:FlAD17.60  
 A:Map position: 3  
 A:introns: 68/1, 101/3, 174/2

Query Match 68.0%; Score 17; DB 2; Length 293;  
 Best Local Similarity 22.2%; Pred. No. 57;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 DB 238 CTADSSSSC 246

RESULT 14  
 T01948



hypothetical protein F1104.13 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 09-Jul-2004  
C/Accession: T01948  
R/Author: Thiele, J.; Stoneking, T.; Langston, Y.; Trevisan, E.  
submitted to the EMBL Data Library, October 1998  
A/Description: The sequence of A. thaliana F1104.  
A/Reference number: Z14466  
A/Accession: T01948  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-298 <ABU>  
A/Cross-references: UNIPROT:O82599; UNIPARC:UP100000A5CDC; EMBL:AF096370; NID:93695372;  
A/Experimental source: cultivar Columbia  
C/Genetics:  
A/Map position: 4  
A/Intons: 158/3  
A/Note: F1104.13  
C/Superfamily: Arabidopsis thaliana hypothetical protein F21F14.120

Query Match 68.0%; Score 17; DB 2; Length 298;  
Best Local Similarity 22.2%; Pred. No. 57;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
DB 200 CXXXXXXXC 208

RESULT 15  
A85019  
probable bHLH DNA-binding protein [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: A85019  
R/Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
Nature 402, 769-777, 1999  
A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A/Reference number: A85001; MUID:20083488; PMID:10617198  
A/Accession: A85019  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-315 <STO>  
A/Cross-references: UNIPROT:Q9M128; UNIPARC:UP1000009DFA; GB:NC\_001268; NID:97268189; F  
C/Genetics:  
A/Map position: 4  
A/Intons: 158/3  
C/Superfamily: Arabidopsis thaliana hypothetical protein F21F14.120

Query Match 68.0%; Score 17; DB 2; Length 315;  
Best Local Similarity 22.2%; Pred. No. 58;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
DB 188 CXXXXXXXC 196

RESULT 16  
AB1224  
cobalamin biosynthesis protein Cbid homolog cbid [imported] - Listeria monocytogenes (str  
C/Species: Listeria monocytogenes  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AB1224  
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fehrl, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitouram, A.; Ma  
Ok, C.; Schleuter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A/Title: Comparative genomics of Listeria species.  
A/Reference number: AB1077; MUID:21537279; PMID:11679669  
A/Accession: AB1224

A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-373 <GLA>  
A/Cross-references: UNIPROT:Q9Y757; UNIPARC:UP10000054CCA; GB:NC\_003210; PIDN:CAC99272.1  
C/Genetics:  
A/Experimental source: strain EGD-e  
A/Map position: 5  
C/Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 68.0%; Score 17; DB 2; Length 373;  
Best Local Similarity 22.2%; Pred. No. 61;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
DB 21 CAAAAKAC 29

RESULT 17  
AD1577  
cobalamin biosynthesis protein Cbid homolog cbid [imported] - Listeria innocua (strain  
C/Species: Listeria innocua  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AD1577  
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fehrl, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitouram, A.; Ma  
Ok, C.; Schleuter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A/Title: Comparative genomics of Listeria species.  
A/Reference number: AB1077; MUID:21537279; PMID:11679669  
A/Accession: AD1577  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-373 <GLA>  
A/Cross-references: UNIPROT:Q92C14; UNIPARC:UP100000CC4A4; GB:AL592022; PIDN:CMC96388.1  
A/Experimental source: strain C1p11262  
C/Genetics:  
A/Map position: 5  
C/Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 68.0%; Score 17; DB 2; Length 373;  
Best Local Similarity 22.2%; Pred. No. 61;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
DB 21 CAAAAKAC 29

RESULT 18  
T43352  
nuclear receptor NHR-13 - Caenorhabditis elegans (fragment)  
C/Species: Caenorhabditis elegans  
C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C/Accession: T43352  
R/Sluder, A.E.; Mathews, S.W.; Hough, D.; Yin, V.P.; Maina, C.V.  
Genome Res. 9, 103-120, 1999  
A/Title: The nuclear receptor superfamily has undergone extensive proliferation and dive  
A/Reference number: Z22443; MUID:99148134; PMID:10022975  
A/Accession: T43352  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-415 <SLU>  
A/Cross-references: UNIPROT:Q9XYB8; UNIPARC:UP100000755B6; EMBL:AF083228; NID:94139081;  
C/Genetics:  
A/Map position: 5  
A/Intons: 158/3  
A/Note: nhr-13  
Query Match 68.0%; Score 17; DB 2; Length 415;  
Best Local Similarity 22.2%; Pred. No. 63;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
Db 12 CSSSSNSC 20

## RESULT 19

T33024 hypothetical protein K07H8.2 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T33024

R/Fulton, B.; Hawkins, J.; Gattung, S.; Wohldmann, P.; Elliott, G.

A/Description: The sequence of C. elegans cosmid K07H8.

A/Reference number: Z21264

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Residues: 1-507 <FUL>

A/Molecule type: DNA

A/Cross-references: UNIPROT:O45182; UNIPARC:UPI0000080E7E; EMBL:AF047659; P1DN:AAC04427.

A/Experimental source: strain Bristol N2; clone K07H8

C/Genetics:

A:Gene: CESP:K07H8.2

A/Map position: 4

A/Introns: 20/3; 55/2; 93/3; 152/3; 447/3

Query Match 68.0%; Score 17; DB 2; Length 507;

Best Local Similarity 22.2%; Pred. No. 66;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 210 CASSLATAC 218

## RESULT 20

T48004 multifunctional aminoacyl-tRNA ligase-like protein - Arabidopsis thaliana

N/Alternate names: protein T17J13.80

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C/Accession: T48004

R/Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, W.; Mewes, H.W.; Lemcke, K.; Mayer, submitted to the Protein Sequence Database, February 2000

A/Reference number: Z24482

A/Accession: T48004

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-530 <RIE>

A/Cross-references: UNIPROT:Q9M1R2; UNIPARC:UPI000009CD43; EMBL:AL138651

A/Experimental source: cultivar Columbia; BAC clone T17J13

C/Genetics:

A/Map position: 3

A/Introns: 64/3; 94/3; 133/3; 196/2; 218/3; 259/3; 310/3; 405/3; 478/3; 510/1

A/Note: T17J13.80

C/Superfamily: proline-tRNA ligase pros

Query Match 68.0%; Score 17; DB 2; Length 530;

Best Local Similarity 22.2%; Pred. No. 67;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 354 CTATASALC 362

## RESULT 21

A42143

C/Species: Caenorhabditis elegans

C/Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C/Accession: A42143; T29535

R/Bowerman, B.; Eaton, B.A.; Priess, J.R.

A/Title: skn-1, a maternally expressed gene required to specify the fate of ventral bias

A/Reference number: A42143; MUID:92191285; PMID:1547503

A/Accession: A42143

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-533 <BOW>

A/Cross-references: UNIPROT:O8MPW3; UNIPARC:UPI0000081925

A/Note: sequence extracted from NCBI backbone (NCBIN:88973, NCBI:88974)

R/Pauley, A.; Gattung, S.

A/Description: The sequence of C. elegans cosmid T19E7.

A/Reference number: Z20637

A/Accession: T29535

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-386 <PAU>

A/Cross-references: UNIPARC:UPI000017A521; EMBL:U42843; P1DN:AAA83594.1

C/Genetics:

A:Gene: CESP:skn-1

A/Introns: 141/3; 249/1; 300/1

C/Keywords: nucleus

F/495-532/Domain: Ios/Jun DNA-binding domain homology <RUD>

Query Match 68.0%; Score 17; DB 2; Length 533;

Best Local Similarity 22.2%; Pred. No. 67;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 384 CTTSSSSTC 392

## RESULT 22

T07739

probable ferredoxinase (EC 4.99.1.1) - potato

C/Species: Solanum tuberosum (potato)

C/Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004

C/Accession: T07739

R/Johnston, D.J.

A/Submitted to the EMBL Data Library, April 1998

A/Reference number: Z15932

A/Accession: T07739

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-543 <JOH>

A/Cross-references: UNIPROT:O64391; UNIPARC:UPI00000A543B; EMBL:AJ005802; P1DN:CAA06705

A/Experimental source: cv. Bintje

C/Function:

A/Description: catalyzes the insertion of iron into protoporphyrin to produce heme

A/Pathway: heme biosynthesis

C/Superfamily: ferredoxinase

C/Keywords: iron; lyase; mitochondrial inner membrane

Query Match 68.0%; Score 17; DB 2; Length 543;

Best Local Similarity 22.2%; Pred. No. 67;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 31 CTTSSFASC 39

## RESULT 23

THH0B

thrombomodulin precursor [validated] - human

C/Species: Homo sapiens (man)

C/Date: 31-Dec-1988 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004

C/Accession: A41442; A28307; A29680; A27073; JX0264; S38954

R/Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maruyama

U. Biochem. 103, 281-285, 1988

A/Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed activ

A:Reference number: A41442; MUID:88227901; PMID:2836377  
A:Accession: A41442  
A:Molecule type: DNA  
A:Residues: 1-575 <SUZ>  
A:Cross-references: UNIPROT:P07204; UNIPARC:UPI00000498EB; DDBJ:D00210; NID:9220126; PID:R:Jackman, R.W.; Beiler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.  
Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987  
A:Title: Human thrombomodulin gene is intron depleted; nucleic acid sequences of the cDN  
A:Reference number: A28307; MUID:87317665; PMID:2819876  
A:Accession: A28307  
A:Molecule type: DNA; mRNA  
A:Residues: 1-472; 'A', 474-575 <JAC>  
A:Cross-references: UNIPARC:UPI00000002BD; GB:J02973; NID:9339658; PIDN:AAA61175.1; PID:R:Suzuki, K.; Kusumoto, H.; Deyashiki, Y.; Nishioaka, J.; Maruyama, I.; Zushi, M.; Kawahara  
EMBO J. 6, 1891-1897, 1987  
A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on endoth  
A:Reference number: A29680; MUID:88004395; PMID:2820710  
A:Accession: A29680  
A:Molecule type: mRNA  
A:Residues: 1-575 <SUZ>  
A:Cross-references: UNIPARC:UPI00000498PB; GB:X05495; NID:937123; PIDN:CAA29045.1; PID:G  
A:Experimental source: lung endothelium  
A:Note: Part of this sequence, including the amino end of the mature protein, were deter  
R.Wen, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.  
Biochemistry 26, 4350-4357, 1987  
A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization of the  
A:Reference number: A27073; MUID:88024950; PMID:2822087  
A:Accession: A27073  
A:Molecule type: mRNA  
A:Residues: 1-472; 'A', 474-575 <WEN>  
A:Cross-references: UNIPARC:UPI00000002BD; GB:M16552; NID:9339656; PIDN:AAB59508.1; PID:  
A:Experimental source: placenta  
A:Note: Parts of this sequence were determined by protein sequencing  
R.Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Ohkuchi, M.; Kimura, S.; Aoki, N.  
J. Biochem. 113, 433-440, 1993  
A:Title: Urinary thrombomodulin, its isolation and characterization.  
A:Reference number: JX0264; MUID:93293792; PMID:8390446  
A:Accession: JX0264  
A:Molecule type: protein; mRNA  
A:Residues: 19-472; 'A', 474-486 <YAM>  
A:Cross-references: UNIPARC:UPI0000173341  
A:Experimental source: urine  
A:Note: the urinary form appears to be identical with that circulating in plasma  
R.Gerlitz, B.; Hassell, T.; Vlahos, C.U.; Parkinson, J.F.; Bang, N.U.; Grinnell, B.W.  
Biochem. J. 295, 131-140, 1993  
A:Title: Identification of the predominant glycosaminoglycan-attachment site in soluble  
serine.  
A:Reference number: S38954; MUID:94029900; PMID:8216207  
A:Accession: S38954  
A:Molecule type: protein  
A:Residues: 475-491; 'X', 493-494 <GER>  
A:Cross-references: UNIPARC:UPI0000173342  
A:Note: the residue designated 'X' was determined to be a Ser with covalently bound chon  
R.Melinger, D.P.; Komives, E.A.  
submitted to the Brookhaven Protein Data Bank, September 1995  
A:Reference number: A67369; PDB:1ZAO  
A:Contents: annotation, conformation and disulfide bond assignments by (1)H-NMR, residue  
R.Tulinsky, A.; Mathews, I.I.  
submitted to the Brookhaven Protein Data Bank, August 1994  
A:Reference number: A52804; PDB:1HLT  
A:Contents: annotation, X-ray crystallography, 3.0 angstroms, residues 426-442  
R.Hirabai, R.; Komives, E.A.; Ni, F.  
submitted to the Brookhaven Protein Data Bank, November 1995  
A:Reference number: A65583; PDB:1FGD  
A:Contents: annotation, conformation by (1)H-NMR, residues 427-444  
Protein Sci. 5, 195-203, 1996  
A:Title: Structural resiliency of an EGF-like subdomain bound to its target protein, the  
A:Reference number: A58595; MUID:96276211; PMID:8745536  
A:Contents: annotation, conformation by (1)H-NMR  
C:Genetics:  
A:Gene: GDB:THBD  
A:Cross-references: GDB:119613; OMIM:188040

A:Map position: 20p11.2-20p11.2  
A:introns: #status absent  
C:Complex: homodimer, urinary form  
C:Function:  
A:Description: inhibits thrombin activation of fibrinogen; cofactor for thrombin activat  
A:Pathway: blood coagulation moderation  
A:Note: the membrane-bound form is located on the endothelium luminal surface of arterie  
A:Note: thrombin complexed with the membrane-bound form is subject to endocytosis  
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
C:Keywords: anticoagulant; beta-hydroxyaspartate; beta-hydroxyaspartic acid; blood coag  
e protein  
F:18/Domain: signal sequence #status predicted <SIG>  
F:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>  
F:19-513/Domain: extracellular #status predicted <EXT>  
F:19-486/Product: thrombomodulin, urinary form #status experimental <MAU>  
F:24-167/Domain: C-type lectin homology <LCH>  
F:177-199/Region: PEPT sequence  
F:201-233/Region: PEPT sequence  
F:245-280/Domain: EGF homology <EG1>  
F:288-323/Domain: EGF homology <EG2>  
F:329-362/Domain: EGF homology <EG3>  
F:369-404/Domain: EGF homology <EG4>  
F:408-439/Domain: EGF homology <EG5>  
F:445-480/Domain: EGF homology <EG6>  
F:485-513/Region: PEPT sequence  
F:517-539/Domain: intracellular #status predicted <INT>  
F:540-575/Domain: transmembrane #status predicted <TMN>  
F:47,115,116,382,409/Binding site: carboxydrate (Asn) (covalent) #status predicted  
F:174,225,411,504/Binding site: carboxydrate (Thr) (covalent) #status predicted  
F:245-256,252-265,267-280,286-296,292-308,310-323,329-340,336-349,351-362,369-378,374-38  
F:334,498/Binding site: carboxydrate (Ser) (covalent) #status predicted  
F:342/Modified site: erythro-beta-hydroxyaspartate (Asn) #status experimental  
F:90,492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimen

Query Match 68.0%; Score 17; DB 1; Length 575;  
Best Local Similarity 22.2%; Pred. No. 69;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Db 280 CTRASTQC 288

RESULT 24  
T07953  
Lectin-like protein zsp2, zygote-specific - Chlamydomonas reinhardtii  
C:Species: Chlamydomonas reinhardtii  
C:date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 09-Jul-2004  
C:Accession: T07953  
R:Suzuki, L.; Woessner, J.P.; Uchida, H.; Kuriyama, H.; Takano, H.; Kuroiwa, H.; Yuasa,  
submitted to the EMBL Data Library, March 1998  
A:Description: A lectin-like protein mediates the assembly of the extracellular matrix c  
A:Reference number: Z16243  
A:Accession: T07953  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-582 <SUZ>  
A:Cross-references: UNIPROT:O65153; UNIPARC:UPI00000454CF; EMBL:AF053099; NID:92997677;  
C:Genetics:  
A:Gene: zsp2  
A:introns: 27/1; 292/3

Query Match 68.0%; Score 17; DB 2; Length 582;  
Best Local Similarity 22.2%; Pred. No. 69;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Db 257 CARSSTTC 265

RESULT 25  
T07952

lectin-like protein, zygote-specific - Chlamydomonas reinhardtii  
 C:Species: Chlamydomonas reinhardtii  
 C:Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 09-Jul-2004  
 C:Accession: T07952  
 R:Suzuki, L.; Woessner, J.P.; Uchida, H.; Kuriyama, H.; Takano, H.; Kuroiwa, H.; Yuasa, submitted to the EMBL Data Library, March 1998  
 A:Description: A lectin-like protein mediates the assembly of the extracellular matrix  
 A:Reference number: Z16243  
 A:Accession: T07952  
 A:Status: preliminary; translated from GB/EMBL/DDBL  
 A:Molecule type: DNA  
 A:Residues: 1-582 <SU2>  
 A:Cross-references: UNIPROT:Q9SB11; UNIPARC:UPI00000454C; EMBL:AF053098; NID:g2997675;  
 C:Genetics:  
 A:Gene: z9p2  
 A:Introns: 27/1, 292/3

Query Match 68.0%; Score 17; DB 2; Length 582;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
 DB 257 CARSTTTC 265

RESULT 26  
 B25682  
 homeotic protein Engrailed - fruit fly (Drosophila virilis)  
 C:Species: Drosophila virilis  
 C:Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 09-Jul-2004  
 C:Accession: B25682  
 R:Kasari, J.A.; Poole, S.J.; Wright, D.K.; O'Farrell, P.H.  
 EMBL J. 5, 3583-3589, 1986  
 A:Title: Sequence conservation in the protein coding and intron regions of the engrailed  
 A:Reference number: A91059; MUID:87161768; PMID:2881781  
 A:Accession: B25682  
 A:Molecule type: DNA  
 A:Residues: 1-584 <KAS>  
 A:Cross-references: UNIPROT:P09145; UNIPARC:UPI000012CA14; GB:X04727; NID:g9173; PIDN:CA  
 C:Genetics:  
 A:Gene: en  
 A:Cross-references: FlyBase:FBgn0013111  
 A:Introns: 470/1; 502/3  
 C:Superfamily: engrailed homeotic protein; homeobox homology  
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation  
 F:487-543/Domain: homeobox homology <HOX>

Query Match 68.0%; Score 17; DB 2; Length 584;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
 DB 384 CSASSSSGC 392

RESULT 27  
 S41011  
 hypothetical protein ZK757.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
 C:Accession: S41011  
 R:Thomas, K.  
 submitted to the EMBL Data Library, December 1993  
 A:Reference number: S41011  
 A:Accession: S41011  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-589 <THO>  
 A:Cross-references: UNIPROT:P34679; UNIPARC:UPI000013CC25; EMBL:Z29121; NID:g438366; PID  
 C:Genetics:  
 A:Introns: 19/2; 54/3; 114/3; 146/3; 178/3; 209/1; 253/1; 302/3; 324/3; 352/1; 392/3; 46

C:Keywords: transmembrane protein  
 ,  
 Query Match 68.0%; Score 17; DB 2; Length 589;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
 DB 100 CITSSSTSC 108

RESULT 28  
 A42086  
 CD30 antigen precursor - human  
 N:Alternate names: Ki-1 antigen; nerve growth factor receptor family member CD30  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
 C:Accession: A42086  
 R:Durkop, H.; Latza, U.; Hummel, M.; Eitelbach, F.; Seed, B.; Stein, H.  
 Cell 68, 421-427, 1992  
 A:Title: Molecular cloning and expression of a new member of the nerve growth factor rec  
 A:Reference number: A42086; MUID:92154659; PMID:1310894  
 A:Accession: A42086  
 A:Molecule type: mRNA  
 A:Residues: 1-595 <DUR>  
 A:Cross-references: UNIPROT:P28908; UNIPARC:UPI000000971; GB:M83554; NID:g180095; PIDN:  
 A:Experimental source: HUT-102 cell line  
 A:Note: sequence extracted from NCBI backbone (NCBIN:82088, NCBI:P:82090)  
 C:Genetics:  
 A:Gene: GDB:CD30; DIS166E  
 A:Cross-references: GDB:11547; OMIM:153243  
 A:Map position: 1p36-1p36  
 C:Keywords: glycoprotein; growth factor receptor; transmembrane protein  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-383/Domain: extracellular #status predicted <EXT>  
 F:384-407/Domain: transmembrane #status predicted <TM>  
 F:408-595/Domain: intracellular #status predicted <CT>  
 F:101,276/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 68.0%; Score 17; DB 2; Length 595;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
 DB 289 CATSATNSC 297

RESULT 29  
 D83286  
 hypothetical protein PA2886 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C:Accession: D83286  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; Lam,  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: D83286  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-600 <STO>  
 A:Cross-references: UNIPROT:Q9H2M1; UNIPARC:UPI000005861; GB:AE004714; GB:AE004091; NID  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA2886

Query Match 68.0%; Score 17; DB 2; Length 600;  
 Best Local Similarity 22.2%; Pred. No. 69;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 92 CAALAAAC 100

## RESULT 30

T15408

hypothetical protein C04F6.3 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C&gt;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C:Accession: T15408

R:Nhan, M.

submitted to the EMBL Data Library, December 1995

A:Description: The sequence of C. elegans cosmid C04F6.

A:Reference number: Z18346

A:Accession: T15408

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-617 &lt;NHA&gt;

A:Cross-references: UNIPROT:Q11174; UNIPARC:UPI000060EB1; EMBL:U42835; NID:g1125760; P

C:Genetics:

A:Gene: CRSP:C04F6.3

A:introns: 28/1; 66/2; 504/1

Query Match 68.0%; Score 17; DB 2; Length 617;  
Best Local Similarity 22.2%; Pred. No. 70;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 524 CTTSTASGC 532

## RESULT 31

A36962

laccase (EC 1.10.3.2) precursor - fungus (*Pilobasidium floriforme*) (ATCC 34873)

N:Alternate names: diphenol oxidase

C:Species: *Pilobasidium neoformanis*, *Cryptococcus neoformans*

C&gt;Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 31-Dec-2004

C:Accession: A36962

R:Williamson, P.R.

J. Bacteriol. 176, 656-664, 1994

A:Title: Biochemical and molecular characterization of the diphenol oxidase of *Cryptococcus*

A:Reference number: A36962; MUID:9411944; PMID:8300520

A:Accession: A36962

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-624 &lt;WIL&gt;

A:Cross-references: UNIPROT:Q9UQZ7; UNIPARC:UPI00001751D2; GB:I22866

C:Genetics:

A:Gene: CNLACT1

A:introns: 22/3; 69/3; 117/1; 137/3; 176/2; 204/2; 223/3; 238/3; 371/2; 426/3; 55

C:Keywords: copper; glycoprotein; oxidoreductase

Query Match 68.0%; Score 17; DB 2; Length 624;  
Best Local Similarity 22.2%; Pred. No. 70;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 243 CTATGSSSC 251

## RESULT 32

A41029

integrin beta-8 chain precursor - human

C:Species: *Homo sapiens* (man)

C&gt;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 31-Dec-2004

C:Accession: A41029

R:Moyle, M.; Napier, M.A.; McLean, J.W.

J. Biol. Chem. 266, 19650-19658, 1991

A:Title: Cloning and expression of a divergent integrin subunit beta-8.

A:Reference number: A41029; MUID:92011767; PMID:1918072

A:Accession: A41029  
A:Molecule type: mRNA  
A:Residues: 1-769 <MOV>  
A:Cross-references: UNIPROT:P26012; UNIPARC:UPI000012DA14; GB:M73780; NID:g184520; PID:  
C:Superfamily: integrin, beta subunit; laminin-type EGF-like homology  
C:Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein.

Query Match 68.0%; Score 17; DB 2; Length 769;  
Best Local Similarity 22.2%; Pred. No. 74;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 47 CASSNAAC 55

## RESULT 33

A46140

diacylglycerol kinase (EC 2.7.1.107) - fruit fly (*Drosophila melanogaster*)C:Species: *Drosophila melanogaster*

C&gt;Date: 22-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A46140

R:Masai, I.; Hosoya, T.; Kojima, S.; Hotta, Y.

Proc. Natl. Acad. Sci. U.S.A. 89, 6030-6034, 1992

A:Title: Molecular cloning of a *Drosophila* diacylglycerol kinase gene that is expressed

A:Reference number: A46140; MUID:92335231; PMID:1321433

A:Accession: A46140

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-791 &lt;MAS&gt;

A:Cross-references: UNIPROT:Q01583; UNIPARC:UPI000016BB61; GB:D11120; NID:g217332; PID:

A:Note: sequence extracted from NCBI backbone (NCBIN:108569; NCBI:P108571)

C:Genetics:

A:Gene: FlyBase:Dgk

A:Cross-references: FlyBase:FBgn0004568

C:Superfamily: fruit fly diacylglycerol kinase; calmodulin repeat homology; protein kin

C:Keywords: EF hand; phosphotransferase

Query Match 68.0%; Score 17; DB 2; Length 791;  
Best Local Similarity 22.2%; Pred. No. 75;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 18 CATAAAVAC 26

## RESULT 34

B46140

diacylglycerol kinase (EC 2.7.1.107) - fruit fly (*Drosophila melanogaster*)C:Species: *Drosophila melanogaster*

C&gt;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 29-Sep-1999

C:Accession: B46140

R:Masai, I.; Hosoya, T.; Kojima, S.; Hotta, Y.

Proc. Natl. Acad. Sci. U.S.A. 89, 6030-6034, 1992

A:Title: Molecular cloning of a *Drosophila* diacylglycerol kinase gene that is expressed

A:Reference number: A46140; MUID:92335231; PMID:1321433

A:Accession: B46140

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-796 &lt;MAS&gt;

A:Cross-references: UNIPARC:UPI00001756B6; GB:D11120

C:Genetics:

A:Gene: FlyBase:Dgk

A:Cross-references: FlyBase:FBgn0004568

C:Superfamily: fruit fly diacylglycerol kinase; calmodulin repeat homology; protein kina

C:Keywords: EF hand; phosphotransferase

Query Match 68.0%; Score 17; DB 2; Length 796;  
Best Local Similarity 22.2%; Pred. No. 75;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

DB 18 CATAAAVAC 26

## RESULT 35

integrin beta-5 chain precursor - human  
 C:Species: Homo sapiens (man)  
 C>Date: 14-Jun-1991 #sequence\_revision 14-Jun-1991 #text\_change 31-Dec-2004  
 C/Accession: A38308; A35775; S12534; S11708  
 R/McClean, J.W.; Vestal, D.J.; Chersesh, D.A.; Bodary, S.C.  
 J. Biol. Chem. 265, 17126-17131, 1990  
 A>Title: CDNA sequence of the human integrin beta-5 subunit.  
 A/Reference number: A38308; UID:91009141; PMID:2211615  
 A/Accession: A38308  
 A/Molecule type: mRNA  
 A/Residues: 1-799 <MCL>  
 A/Cross-references: UNIPROT:P18084; UNIPARC:UPI00004D05B; GB:J05633; NID:g186504; PIDN:  
 A/Note: parts of this sequence, including the amino end of the mature protein, were conf  
 R/Suzuki, S.; Huang, Z.S.; Tanihara, H.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 5354-5358, 1990  
 A>Title: Cloning of an integrin beta subunit exhibiting high homology with integrin beta  
 A/Reference number: A35775; UID:90319111; PMID:2371275  
 A/Accession: A35775  
 A/Molecule type: mRNA  
 A/Residues: 1-192 'A', 194-644, 'L', 646-789, 793-799 <SUZ>  
 A/Cross-references: UNIPARC:UPI000016ADB; GB:M35011; NID:g184524; PIDN:AA52707.1; PID:  
 R/Ramaswamy, H.; Hemler, M.E.  
 EMBO J. 9, 1561-1568, 1990  
 A>Title: Cloning, primary structure and properties of a novel human integrin beta subunit  
 A/Reference number: S12534; UID:90228356; PMID:2328726  
 A/Accession: S12534  
 A/Molecule type: mRNA  
 A/Residues: 1-644, 'L', 646-799 <RAM>  
 A/Cross-references: UNIPARC:UPI000012DA10; EMBL:X53002; NID:g33952; PIDN:CAA37188.1; PID:  
 C/Genetics:  
 A/Gene: GDB:ITGB5  
 A/Cross-references: GDB:128005; OMIM:147561  
 A/Map position: 17q11-17qter  
 C/Superfamily: Integrin, beta subunit; laminin-type EGF-like homology  
 C/Keywords: cell adhesion; cytoskeleton; duplication; extracellular matrix; glycoprotein  
 F/1-24/Domain: signal sequence #status predicted <SIG>  
 F/25-799/Product: integrin beta-5 chain #status experimental <MAT>  
 F/25-719/Domain: extracellular #status predicted <EXT>  
 F/463-513/Domain: laminin-type EGF-like homology <LEG>  
 F/720-742/Domain: transmembrane #status predicted <TM>  
 F/743-799/Domain: intracellular #status predicted <INT>  
 F/347,460,477,505,552,586,654,705/Binding site: carbohydrate (asn) (covalent) #status pr

Query Match 68.0%; Score 17; DB 2; Length 799;  
 Best Local Similarity 22.2%; Pred. No. 75;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 28 CTSGSATSC 36

## RESULT 36

T05201  
 hypothetical protein F410.140 - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 C/Accession: T05201  
 R/Bevan, M.; Peters, S.A.; van Staveren, M.; Dirksen, W.; Stiekema, W.; Hohnesiel, J.; New  
 submitted to the Protein Sequence Database, February 1999  
 A/Reference number: Z15402  
 A/Accession: T05201  
 A/Molecule type: DNA  
 A/Residues: 1-934 <BEV>  
 A/Cross-references: UNIPROT:Q9SMY8; UNIPARC:UPI000009D8D1; EMBL:AL035525  
 A/Experimental source: cultivar Columbia; BAC clone F410  
 C/Genetics:

A/Map position: 4  
 A/Introns: 123/2; 203/3; 248/1; 281/2; 347/2; 437/3; 470/3; 500/2; 533/3; 590/3; 641/3;  
 A/Note: F410.140

Query Match 68.0%; Score 17; DB 2; Length 934;  
 Best Local Similarity 22.2%; Pred. No. 79;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 701 CUSATATAC 709

## RESULT 37

T41932  
 hypothetical protein U30 - human herpesvirus 7 (strain J1)  
 C/Species: human herpesvirus 7  
 A/Variety: strain J1  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C/Accession: T41932  
 R/Nicholas, J.  
 submitted to the EMBL Data Library, December 1995  
 A/Description: Determination and analysis of the complete nucleotide sequence of human h  
 A/Reference number: Z22022  
 A/Accession: T41932  
 A/Molecule type: DNA  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Residues: 1-938 <NIC>  
 A/Cross-references: UNIPROT:P52438; UNIPARC:UPI00000007D8; EMBL:U43400; PIDN:AA54692.1  
 A/Experimental source: strain J1  
 C/Genetics:  
 A/Note: U30

Query Match 68.0%; Score 17; DB 2; Length 938;  
 Best Local Similarity 22.2%; Pred. No. 79;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 689 CUSTLATTC 697

## RESULT 38

T15881  
 hypothetical protein D1044.3 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C/Accession: T15881  
 R/Pauley, A.  
 submitted to the EMBL Data Library, June 1994  
 A/Description: The sequence of C. elegans comid D1044.  
 A/Reference number: Z18423  
 A/Accession: T15881  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-1895 <PAU>  
 A/Cross-references: UNIPARC:UPI000004CAE8; EMBL:U00065; NID:g495681; PID:g495684; PIDN:A  
 A/Experimental source: strain Bristol N2  
 C/Genetics:  
 A/Gene: CESP:D1044.3  
 A/Introns: 35/2; 63/2; 150/1; 191/3; 207/2; 405/3; 551/3; 629/1; 764/1; 933/1; 959/2; 10

Query Match 68.0%; Score 17; DB 2; Length 1895;  
 Best Local Similarity 22.2%; Pred. No. 96;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 1233 CUSSSSSSC 1241

## RESULT 39

S50820

surface protein type 51B - Parametium tetraurelia

C/Species: Parametium tetraurelia

C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C/Accession: S50820

R/Scott, J.; Lebeck, C.; Porney, J.

Nucleic Acids Res. 22, 5079-5084, 1994

A/Title: Analysis of the micronuclear B type surface protein gene in Parametium tetraurelia

A/Reference number: S50820; PMID:95098630; PMID:7800503

A/Accession: S50820

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-2395 <SC0>

A/Cross-references: UNIPROT:Q27167; UNIPARC:UPI000007D4F6; EMBL:U07603; NID:9467226; PID

C/Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T31307

R/Zhu, G.; Marchewka, M.J.; Woods, K.M.; Upton, S.J.; Keithly, J.S.

Submitted to the EMBL Data Library, August 1998

A/Description: Characterization of a type I fms gene in the parasitic protozoan Cryptosporidium parvum

A/Reference number: Z20993

A/Accession: T31307

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-8243 <ZHU>

A/Cross-references: UNIPROT:Q96554; UNIPARC:UPI000011010F; EMBL:AF082993; NID:94092068;

C/Genetics:

A/Note: FAS1

C/Keywords: carrier protein

F/100-714/Domain: acetate-CoA ligase homology <ACLI>

F/791-861/Domain: acyl carrier protein homology <ACPI>

F/906-1308/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>

F/2924-2992/Domain: acyl carrier protein homology <ACP2>

F/3062-3463/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>

F/5157-5227/Domain: acyl carrier protein homology <ACP3>

F/5274-5689/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>

Query Match

Best Local Similarity 22.2%; Score 17; DB 2; Length 8243;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

DB 3210 CASLTATC 3218

RESULT 43

G82010

C/Species: Neisseria meningitidis (strain Z2491 serogroup

C/Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004

C/Accession: G82010

R/Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holtz, S.; Jorgensen, K.; Leather, S.; Mouton, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A/Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A/Reference number: AB1775; PMID:20222556; PMID:10761919

A/Accession: G82010

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-52 <PAR>

A/Cross-references: UNIPROT:Q9JWY4; UNIPARC:UPI00000C48B2; GB:AL162752; GB:AL157959; NIT

C/Genetics:

A/Experimental source: serogroup A, strain Z2491

A/Genetic code: SGC5

C/Superfamily: G surface protein

Query Match

Best Local Similarity 22.2%; Score 16; DB 2; Length 52;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

DB 8 CTTSLAAC 16

RESULT 41

A23475

G surface protein - Parametium primaurelia

C/Species: Parametium primaurelia

C/Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 09-Jul-2004

C/Accession: A23475

R/Pratt, A.; Katinka, M.; Garon, F.; Meyer, E.

J. Mol. Biol. 189, 47-60, 1986

A/Title: Nucleotide sequence of the Parametium primaurelia G surface protein. A huge pro

A/Reference number: A23475; PMID:87060934; PMID:3783679

A/Accession: A23475

A/Molecule type: DNA

A/Residues: 1-2718 <PRA>

A/Cross-references: UNIPROT:PI3837; UNIPARC:UPI0000177F7E

A/Note: the authors translated the codon TGC for residue 2665 as TTP

C/Genetics:

A/Genetic code: SGC5

C/Superfamily: G surface protein

RESULT 44  
H66753  
prophage p12 protein 25 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
A:Accession: H66753  
R:Bioclin, A.; Wincker, P.; Mueger, S.; Jallion, O.; Malarre, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A:Reference number: A86625; MUID:21235186; PMID:11337471  
A:Accession: H66753  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-87 <STO>  
A:Cross-references: UNIPROT:Q9CGR4; UNIPARC:UPI000006978; GB:AE005176; PID:g12723983; F  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: p1225

Query Match 64.0%; Score 16; DB 2; Length 87;  
Best Local Similarity 22.2%; Pred. No. 1.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 30 CTAKSSLSC 38

RESULT 45  
JS0036  
Clara cell 10K protein precursor - human  
N:Alternate names: urinary protein 1  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
A:Accession: JS0036; PS0309; A56890; I38397  
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Phillips, S.; Kennedy, A.L.; Anthony, J.; Squeglia  
Biochim. Biophys. Acta 950, 329-337, 1988  
A:Title: Amino-acid and cDNA nucleotide sequences of human Clara cell 10KDa protein.  
A:Reference number: JS0036; MUID:8900784; PMID:3167058  
A:Accession: JS0036  
A:Molecule type: mRNA  
A:Residues: 1-91 <SIN>  
A:Cross-references: UNIPROT:P11684; UNIPARC:UPI00000369D; GB:X13197; NID:g2113; PIDN:C  
A:Accession: PS0309  
A:Molecule type: protein  
A:Residues: 22-23, 'X', 25-28, 'X', 30-31, 'X', 33-36 <SIG>  
A:Cross-references: UNIPARC:UPI0000177C39  
R:Bernard, A.; Reels, H.; Lauwerys, R.; Witters, R.; Gielens, C.; Somillion, A.; Van De  
Clin. Chim. Acta 207, 239-249, 1992  
A:Title: Human urinary protein 1: evidence for identity with the Clara cell protein and  
A:Reference number: A56890; MUID:93009001; PMID:1395029  
A:Accession: A56890  
A:Molecule type: protein  
A:Residues: 22-45 <BBR>  
A:Cross-references: UNIPARC:UPI0000036975  
A:Experimental source: urine  
A:Note: sequence extracted from NCBI backbone (NCBIP:119391)  
R:Hay, J.G.; Daniel, C.; Chu, C.; Crystal, R.G.  
Am. J. Physiol. 268, 565-575, 1995  
A:Title: Human CC10 gene expression in airway epithelium and subchromosomal locus sugges  
A:Reference number: I38397  
A:Accession: I38397  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-91 <RBS>  
A:Cross-references: UNIPARC:UPI00000369D; EMBL:U01101; NID:g457932; PIDN:AAA81885.1; PI  
C:Comment: This protein consists of two identical polypeptides linked by two disulfide b  
C:Genetics:  
A:Gene: CC10  
C:Superfamily: uteroglobin  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-91/Product: Clara cell 10K protein #status experimental <MAT>

Query Match 64.0%; Score 16; DB 2; Length 91;  
Best Local Similarity 22.2%; Pred. No. 1.1e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 16 CSSASSDIC 24

RESULT 46  
UGMS  
uteroglobin precursor - mouse  
N:Alternate names: CC10; Clara cell 10K protein precursor; Clara cell secretory protein  
C:Species: Mus musculus (house mouse)  
C:Date: 03-May-1994 #sequence\_revision 21-Jan-1997 #text\_change 09-Jul-2004  
A:Accession: A53025; A56656; I51925; S24783  
R:Stripp, B.R.; Hultman, J.A.; Bohinski, R.J.  
Genomics 20, 27-35, 1994  
A:Title: Structure and regulation of the murine Clara cell secretory protein gene.  
A:Reference number: A53025; MUID:94292183; PMID:8020953  
A:Accession: A53025  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-96 <STR>  
A:Cross-references: UNIPROT:Q06318; UNIPARC:UPI0000011D0; GB:L24372; NID:g461147; PIDN:  
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Kennedy, A.L.  
Exp. Lung Res. 19, 67-75, 1993  
A:Title: Mouse Clara cell 10-KDa (CC10) protein: cDNA nucleotide sequence and molecular  
A:Reference number: A56656; MUID:93178380; PMID:8440203  
A:Accession: A56656  
A:Molecule type: mRNA; protein  
A:Residues: 1-96 <SIN>  
A:Cross-references: UNIPARC:UPI0000011D0; EMBL:X67702; NID:g49690; PIDN:CAA47936.1; PID  
A:Experimental source: lung  
A:Note: sequence extracted from NCBI backbone (NCBIP:126148)  
A:Note: parts of this sequence, including the amino end of the mature protein, were conf  
R:Marraf, L.R.; Finegold, M.J.; Stanley, L.A.; Major, A.; Watkins, H.K.; DeMayo, F.J.  
Am. J. Respir. Cell Mol. Biol. 9, 231-238, 1993  
A:Title: Cloning and tissue-specific expression of the cDNA for the mouse Clara cell 10K  
A:Reference number: I51925; MUID:94000840; PMID:8398159  
A:Accession: I51925  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-96 <RBS>  
A:Cross-references: UNIPARC:UPI0000011D0; GB:L04503; NID:g20213; PIDN:AAA03625.1; PID  
C:Genetics:  
A:Intons: 19/1; 81/3  
C:Complex: homodimer linked by two disulfide bonds  
C:Superfamily: uteroglobin  
C:Keywords: lung; steroid binding; uterus  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-96/Product: uteroglobin #status predicted <MAT>  
F:24/Disulfide bonds: interchain (to 90) #status predicted  
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 64.0%; Score 16; DB 1; Length 96;  
Best Local Similarity 22.2%; Pred. No. 1.2e+02;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 16 CSSASSDIC 24

RESULT 47  
A36581  
polychlorinated biphenyl-binding protein precursor - rat  
N:Alternate names: Clara cell 10K secretory protein  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 09-Jul-2004  
A:Accession: A36581; S10185; S21676  
R:Nordlund-Moeller, L.; Andersson, O.; Ahlgren, R.; Schilling, J.; Gyllner, M.; Gustafss



J. Biol. Chem. 265, 12690-12693, 1990  
 A:Title: Cloning, structure, and expression of a rat binding protein for polychlorinated  
 A:Reference number: A36581, MUID:9034256; PMID:2115524  
 A:Accession: A36581  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-96 <NOR>  
 A:Cross-references: UNIPROT:P17559; UNIPARC:UPI0000112E8B; GB:J05536; NID:g206039; PIDN:  
 R:Ragen, G.; Wolf, M.; Karyal, S.L.; Singh, G.; Beato, M.; Suske, G.  
 Nucleic Acids Res. 18, 2939-2946, 1990  
 A:Title: Tissue-specific expression, hormonal regulation and 5'-flanking gene region of  
 A:Reference number: S10185, MUID:90272398; PMID:2349092  
 A:Accession: S10185  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-18 <HNG>  
 A:Cross-references: UNIPARC:UPI00001707CC; EMBL:X51318; NID:g55536; PIDN:CAM35701.1; PID  
 R:Umland, T.C.; Swannathan, S.; Furey, W.; Singh, G.; Pletcher, J.; Sax, M.  
 J. Mol. Biol. 224, 441-448, 1992  
 A:Title: Refined structure of rat Clara cell 17 kDa protein at 3.0 Å resolution.  
 A:Reference number: S21676; MUID:92219263; PMID:1560460  
 A:Content: annotation, X-ray crystallography, 3.0 angstroms  
 C:Superfamily: uteroglobin  
 F:24/Disulfide bonds: interchain (to 24) #status experimental  
 F:90/Disulfide bonds: interchain (to 24) #status experimental

Query Match 64.0%; Score 16; DB 2; Length 96;  
 Best Local Similarity 22.2%; Pred. No. 1.2e+02;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 16 CSSASSDIC 24

RESULT 48  
 A10334  
 conserved hypothetical protein YPO2745 [imported] - *Yersinia pestis* (strain CO92)  
 C:Species: *Yersinia pestis*  
 C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: A10334  
 R:Parhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
 deo-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: A10334  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-98 <KUR>  
 A:Cross-references: UNIPROT:Q8ZD47; UNIPARC:UPI00000CD967; GB:AL590842; PIDN:CAC92984.1;  
 C:Genetics:  
 A:Gene: YPO2745  
 C:Superfamily: C4-dicarboxylate carrier protein

Query Match 64.0%; Score 16; DB 2; Length 98;  
 Best Local Similarity 22.2%; Pred. No. 1.2e+02;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 8 CSAETAC 16

RESULT 49  
 S76421  
 ferredoxin [2Fe-2S] - *Synechocystis* sp. (strain PCC 6803)  
 C:Species: *Synechocystis* sp.  
 A:Vareity: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
 C:Accession: S76421  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.

O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
 DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S76421  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-103 <KAN>  
 A:Cross-references: UNIPROT:P74449; UNIPARC:UPI00000D7139; EMBL:D90915; GB:AB001339; NID  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 C:Keywords: 2Fe-2S; iron-sulfur protein; metalloprotein  
 F:28-84/Domain: ferredoxin [2Fe-2S] homology <FER>  
 F:43,48,51,83/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 64.0%; Score 16; DB 2; Length 103;  
 Best Local Similarity 22.2%; Pred. No. 1.2e+02;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 83 CAAYATSDC 91

RESULT 50  
 J00863  
 hypothetical 11.6k protein - *Escherichia coli* retron Ec67  
 C:Species: *Escherichia coli* retron Ec67  
 C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 09-Jul-2004  
 C:Accession: J00863  
 R:Hsu, M.Y.; Inouye, M.; Inouye, S.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 9454-9458, 1990  
 A:Title: Retron for the 67-base multicopy single-stranded DNA from *Escherichia coli*: a p  
 A:Reference number: J00851; MUID:91067724; PMID:1701261  
 A:Accession: J00863  
 A:Molecule type: DNA  
 A:Residues: 1-104 <HSU>  
 A:Cross-references: UNIPROT:P21322; UNIPARC:UPI000013BF85; GB:M55249; NID:g145143; PIDN:  
 A:Experimental source: *E. coli* strain Cl-1  
 C:Genetics:  
 A:Note: insertion site is equivalent to 19 min of *E. coli* K12 genetic map  
 C:Superfamily: *Escherichia coli* retron Ec67 hypothetical 11.6k protein

Query Match 64.0%; Score 16; DB 2; Length 104;  
 Best Local Similarity 22.2%; Pred. No. 1.2e+02;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXC 9  
 DB 5 CSAESAHSC 13

Search completed: January 4, 2006, 16:10:27  
 Job time: 10.313 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 58.7739 Seconds  
(without alignments)  
108.037 Million cell updates/sec

Title: US-09-932-322-10  
Perfect score: 25  
Sequence: 1 CXXXXXXC 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : Uniprot\_05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	72.0	144	2	061H6_drosophila
2	18	72.0	155	2	P91214_CABEL
3	18	72.0	269	2	07PR07_anopheles
4	18	72.0	281	2	08G050_MAIZE
5	18	72.0	281	2	08LXK0_MAIZE
6	18	72.0	347	2	08G050_9BBTA
7	18	72.0	416	2	09ST87_ORYSA
8	18	72.0	468	2	07XST8_ORYSA
9	18	72.0	470	2	09LFB4_ARATH
10	18	72.0	504	2	09XZ8_LEIMA
11	18	72.0	510	2	08S4F6_ARATH
12	18	72.0	513	2	07XZ21_XENIA
13	18	72.0	518	2	0640H3_XENIA
14	18	72.0	569	2	07OKT3_GIALA
15	18	72.0	684	2	07XV79_ORYSA
16	18	72.0	772	1	DILA_BRARE
17	18	72.0	828	2	04WY29_ASPTU
18	18	72.0	1013	2	04SR96_TETNG
19	18	72.0	1031	2	04QFB5_LEIMA
20	18	72.0	1107	2	05CV70_CRYPV
21	18	72.0	1199	1	E75BC_DROME
22	18	72.0	1355	1	E75BA_DROME
23	18	72.0	1365	1	09VU79_DROME
24	18	72.0	1412	1	E75BB_DROME
25	18	72.0	1793	2	09W451_DROME
26	18	72.0	1836	2	08MP07_DROME
27	18	72.0	1894	2	04OS15_LEIMA
28	18	72.0	2087	2	04S488_TETNG
29	17	68.0	50	2	09L981_VIRCH
30	17	68.0	71	2	04RA29_TETNG
31	17	68.0	96	2	08VD96_MESAU

32	17	68.0	108	2	04P4H6_USTMA	04P4H6 usiliago ma
33	17	68.0	108	2	09ZSP6_ARATH	09ZSP6 arabidopsis
34	17	68.0	115	2	07ZD14_DESVH	07ZD14 desulevibr
35	17	68.0	117	2	06VZRI_CNPV	06VZRI canarypox v
36	17	68.0	119	2	04TWL2_ASP	04TWL2 african swi
37	17	68.0	119	2	04TWL3_ASP	04TWL3 african swi
38	17	68.0	119	2	04TWL5_ASP	04TWL5 african swi
39	17	68.0	119	2	04TWL6_ASP	04TWL6 african swi
40	17	68.0	119	2	04TWL7_ASP	04TWL7 african swi
41	17	68.0	119	2	04TWL8_ASP	04TWL8 african swi
42	17	68.0	119	2	04TWL9_ASP	04TWL9 african swi
43	17	68.0	125	2	04TWM0_ASP	04TWM0 african swi
44	17	68.0	125	2	05YPN5_NOCFA	05YPN5 nocardia fa
45	17	68.0	133	2	04RTU4_TETNG	04RTU4 tetradon n
46	17	68.0	134	2	06ZRF5_HUMAN	06ZRF5 homo sapien
47	17	68.0	147	2	09CMW3_MOUSE	09CMW3 mus musculu
48	17	68.0	155	2	06F340_ORYSA	06F340 oryza sativ
49	17	68.0	155	2	04TWL4_ASP	04TWL4 african swi
50	17	68.0	157	1	C1025_MOUSE	04TWL50 homo sapien
51	17	68.0	157	1	C1025_MOUSE	094772 mus musculu
52	17	68.0	157	2	04Q6M1_LEIMA	04Q6M1 leishmania
53	17	68.0	160	2	04T116_AZOVI	04T116 azotobacter
54	17	68.0	163	2	09BIMO_GIALA	09BIMO giardia lam
55	17	68.0	166	2	05TS94_HUMAN	05TS94 homo sapien
56	17	68.0	168	2	06K5J0_ORYSA	06K5J0 oryza sativ
57	17	68.0	170	2	08PRC3_XANAC	08PRC3 xanthomonas
58	17	68.0	171	2	04L1S2_YENTR	04L1S2 salmonella
59	17	68.0	171	2	093GN6_SALTY	093GN6 salmonella
60	17	68.0	172	2	05TS90_HUMAN	05TS90 homo sapien
61	17	68.0	172	2	06K2N4_ORYSA	06K2N4 oryza sativ
62	17	68.0	175	2	09VSQ7_DROME	09VSQ7 drosophila
63	17	68.0	183	2	05TS95_HUMAN	05TS95 homo sapien
64	17	68.0	186	1	C1025_BRABR	06G660 brachydanio
65	17	68.0	191	2	0502H7_BRABR	0502H7 brachydanio
66	17	68.0	199	2	076588_CABBL	076588 caenorhabdi
67	17	68.0	202	2	06EX59_CANGA	06EX59 candida gla
68	17	68.0	204	2	06YU14_ORYSA	06YU14 oryza sativ
69	17	68.0	207	2	08N1Y5_HUMAN	08N1Y5 homo sapien
70	17	68.0	212	1	YN87_YEAST	P53719 saccharomyc
71	17	68.0	214	2	053846_MYCTU	053846 mycobacteri
72	17	68.0	214	2	07U1S8_MYCTU	07U1S8 mycobacteri
73	17	68.0	222	2	04V5Y8_DROME	04V5Y8 drosophila
74	17	68.0	225	2	09LJBT_ARATH	09LJBT arabidopsis
75	17	68.0	228	2	04V5S4_DROME	04V5S4 drosophila
76	17	68.0	244	2	09Y9Q2_AERYE	09Y9Q2 aeoropyrum p
77	17	68.0	247	2	06BP27_DEBDA	06BP27 debaryomyc
78	17	68.0	248	2	09US89_SCHPO	09US89 echinosach
79	17	68.0	254	2	06WMS9_BRABR	06WMS9 brachiolesto
80	17	68.0	257	2	06ZD11_ORYSA	06ZD11 oryza sativ
81	17	68.0	259	2	083212_TREPFA	083212 treponema p
82	17	68.0	262	1	W0X3A_MAIZE	07U0V1 zea mays (m
83	17	68.0	265	1	W0X3B_MAIZE	06S313 zea mays (m
84	17	68.0	273	2	07SH84_ORYSA	07SH84 oryza sativ
85	17	68.0	281	2	057N21_SALGI	057N21 salmonella
86	17	68.0	281	2	05HQJ3_SALPIA	05HQJ3 salmonella
87	17	68.0	281	2	08Z796_SALTY	08Z796 salmonella
88	17	68.0	281	2	08ZP77_SALTY	08ZP77 salmonella
89	17	68.0	283	2	07D973_MYCTU	07D973 mycobacteri
90	17	68.0	286	2	06G0D6_XENIA	06G0D6 xenopus lae
91	17	68.0	290	2	06A411_PROAC	06A411 propionibac
92	17	68.0	291	2	08W318_PRODI	08W318 vitis labru
93	17	68.0	293	1	XTH31_ARATH	P93046 arabidopsis
94	17	68.0	298	2	082599_ARATH	082599 arabidopsis
95	17	68.0	305	2	04LMH9_YRUTK	04LMH9 burkholderi
96	17	68.0	305	2	06TL98_SYWTH	06TL98 symbiodacte
97	17	68.0	306	2	07F8S0_ORYSA	07F8S0 oryza sativ
98	17	68.0	306	2	06ZJ81_BURMA	06ZJ81 burkholderi
99	17	68.0	308	2	06GM13_CABEL	06GM13 caenorhabdi
100	17	68.0	309	2	08GY34_ARATH	08GY34 arabidopsis
101	17	68.0	309	2	04NURS_9DELT	04NURS anaeromyxob
102	17	68.0	310	2	05N9V7_ORYSA	05N9V7 oryza sativ
103	17	68.0	311	2	06EP90_ORYSA	06EP90 oryza sativ
104	17	68.0	311	2	0635X3_BURPS	0635X3 burkholderi

105	17	68.0	315	2	057TV3_9TRYP	057V3 trypanosoma	178	17	68.0	602	2	065169_ASE	065169 african swi
106	17	68.0	315	2	09M128_ARATH	09m128 arbidopsis	179	17	68.0	617	1	CHIT_CAEEL	011174 caenorhabdi
107	17	68.0	319	2	04YRS3_PLABE	04yrs3 plasmodium	180	17	68.0	617	1	0619N2_CAEBR	0619n2 caenorhabdi
108	17	68.0	322	2	09FER3_ORYSA	09ftrg3 oryza sativ	181	17	68.0	623	1	SKN1_CAEEL	013707 caenorhabdi
109	17	68.0	324	1	YO41_CAEEL	P34679 caenorhabdi	182	17	68.0	623	2	04LK90_9EURK	04lk90 burkholderi
110	17	68.0	326	2	089G34_BRAJA	089g34 bradyrhizob	183	17	68.0	624	2	05KEA0_CRYNE	05kea0 cryptococcu
111	17	68.0	336	2	06S7A3_ARATH	06s7a3 arabidopsis	184	17	68.0	624	2	05SP57_CRYNE	05sp57 cryptococcu
112	17	68.0	336	2	09L1A6_ARATH	09l1a6 arbidopsis	185	17	68.0	625	2	04P2X2_USTMA	04p2x2 ustilago ma
113	17	68.0	338	2	062AK4_BURMA	062ak4 burkholderi	186	17	68.0	627	2	06FNF7_CANGA	06fnf7 candida gla
114	17	68.0	344	2	08T004_METYA	08t004 methanopyru	187	17	68.0	629	2	06EAM6_PRA	06eam6 piuum sativ
115	17	68.0	345	2	09DW86_RCWMV	09dw86 rat ctyomeg	188	17	68.0	633	2	08IHC0_DROME	08ihc0 drosophila
116	17	68.0	353	2	05EXJ3_9LAMI	05exj3 streptocarp	189	17	68.0	633	2	0818W0_GIALA	0818w0 giardia lam
117	17	68.0	373	1	CBID_LISIN	092c14 listeria in	190	17	68.0	638	2	07QOC4_GIALA	07qoc4 giardia lam
118	17	68.0	373	1	CBID_LISMF	0720n5 listeria mo	191	17	68.0	663	2	096UAB_NEUCR	096uab neurospora
119	17	68.0	373	1	CBID_LISMO	08y747 listeria mo	192	17	68.0	667	2	0883X3_PSSSM	0883x3 pseudomonas
120	17	68.0	380	2	061U40_CAEBR	061u40 caenorhabdi	193	17	68.0	672	1	PON25_CAEEL	01951 caenorhabdi
121	17	68.0	387	2	05KE99_CRYNE	05ke99 cryptococcu	194	17	68.0	674	2	07S7N5_NEUCR	07s7n5 neurospora
122	17	68.0	388	2	040920_PICMA	040920 picea maria	195	17	68.0	687	2	0948Y7_VOLCA	0948y7 volvox cart
123	17	68.0	388	2	09CK39_PASMU	09ck39 pasteurella	196	17	68.0	705	2	0818W1_GIALA	0818w1 giardia lam
124	17	68.0	394	2	04T7W8_TERNG	04t7w8 tetraodon n	197	17	68.0	710	2	04NV42_9DELT	04nv42 anaeromyxob
125	17	68.0	402	2	075NB1_PLBOS	075nb1 pleurotus o	198	17	68.0	731	2	07ZXU8_XENIA	07zxu8 xenopus lae
126	17	68.0	402	2	075PR3_9AGAR	075pr3 pleurotus e	199	17	68.0	733	2	06L8T9_CUCSA	06l8t9 cucumis sat
127	17	68.0	402	2	075VR2_9AGAR	075vr2 pleurotus e	200	17	68.0	742	2	04VPP9_9BURK	04vpp9 burkholderi
128	17	68.0	404	2	05QUE1_IDILLO	05que1 idiomarina	201	17	68.0	751	2	09NK92_DROME	09nk92 drosophila
129	17	68.0	408	2	08J2S3_9HOMO	08j2s3 hericium er	202	17	68.0	752	2	06FEZ6_ACTAD	06fez6 actinetobact
130	17	68.0	414	1	NHR13_CAEEL	09n4g7 caenorhabdi	203	17	68.0	768	2	09VC61_DROME	09vc61 drosophila
131	17	68.0	415	2	09XYB8_CAEEL	09xyb8 caenorhabdi	204	17	68.0	768	2	0505D9_MOUSE	0505d9 mus musculu
132	17	68.0	415	2	06F3C8_CAEEL	06f3c8 caenorhabdi	205	17	68.0	769	1	ITB8_HUMAN	01583 drosophila
133	17	68.0	425	2	082P91_STRAM	082p91 streptomyce	206	17	68.0	791	1	DKT1_DROME	001583 drosophila
134	17	68.0	427	2	072869_LEEND	072869 lentinula e	207	17	68.0	791	2	051YF9_DROME	051yf9 drosophila
135	17	68.0	427	2	05W612_CAEEL	05w612 caenorhabdi	208	17	68.0	791	2	09GK49_BOVIN	09gk49 bos taurus
136	17	68.0	433	2	06IR40_MOUSE	06ir40 mus musculu	209	17	68.0	798	1	ITB5_MOUSE	070309 mus musculu
137	17	68.0	439	2	09BUC9_HUMAN	09buc9 homo sapien	210	17	68.0	799	2	ITB5_HUMAN	06pe70 mus musculu
138	17	68.0	457	2	062GPR_BURMA	062gpr burkholderi	211	17	68.0	799	2	06PE70_MOUSE	06pe70 mus musculu
139	17	68.0	464	2	06DGV8_BRABE	06dgv8 brachydanto	212	17	68.0	800	2	08SQB9_BOVIN	08sqb9 bos taurus
140	17	68.0	467	2	06JXP7_BURPS	06jxp7 burkholderi	213	17	68.0	807	2	07RG33_PLAYO	07rg33 plasmodium
141	17	68.0	468	2	05YC24_ARATH	05yc24 arabidopsis	214	17	68.0	812	2	06T6B3_CHICK	06t6b3 gallus galli
142	17	68.0	470	2	09FILT_ARATH	09fil7 arbidopsis	215	17	68.0	831	2	04T505_TERNG	04t505 tetraodon n
143	17	68.0	475	2	04SU64_TERNG	04su64 tetraodon n	216	17	68.0	858	2	07R6P7_GIALA	07r6p7 giardia lam
144	17	68.0	484	2	05SP58_CRYNE	05sp58 cryptococcu	217	17	68.0	858	2	09VCM2_DROME	09vcm2 drosophila
145	17	68.0	484	2	09VSE9_DROME	09ves9 drosophila	218	17	68.0	884	2	04YRS2_TERNG	04yrs2 tetraodon n
146	17	68.0	490	2	08BO60_MOUSE	08bq60 mus musculu	219	17	68.0	885	2	04IFB9_GIBZE	04ifb9 gibberella
147	17	68.0	494	2	08MS79_DROME	08ms79 drosophila	220	17	68.0	890	2	09V435_DROME	09v435 drosophila
148	17	68.0	503	2	096SK0_CAEEL	096sk0 caenorhabdi	221	17	68.0	903	2	04FXR3_LEIMA	04fxr3 leishmania
149	17	68.0	505	2	0624P3_CABBR	0624p3 caenorhabdi	222	17	68.0	907	2	04P9R6_USTMA	04p9r6 ustilago ma
150	17	68.0	507	2	04S182_CAEEL	04s182 caenorhabdi	223	17	68.0	907	2	05TAX2_HUMAN	05tax2 homo sapien
151	17	68.0	507	2	08S2T4_DROME	08s2t4 drosophila	224	17	68.0	909	2	04QGS5_LEIMA	04qgs5 leishmania
152	17	68.0	510	2	057TUG_9TRYP	057tug trypanosoma	225	17	68.0	934	2	09SMY8_ARATH	09smy8 arabidopsis
153	17	68.0	519	2	04QGN3_LEIMA	04qgn3 leishmania	226	17	68.0	938	1	V120_HHVU	09ey88 human herpe
154	17	68.0	522	2	086179_DICDI	086179 dictyostell	227	17	68.0	938	2	07YV77_9BETA	07y777 human herpe
155	17	68.0	522	2	0571A7_MOUSE	0571a7 mus musculu	228	17	68.0	972	1	DCR1A_CHICK	05q1c4 gallus galli
156	17	68.0	530	2	09M1R2_ARATH	09m1r2 arabidopsis	229	17	68.0	980	2	09FHD0_ARATH	09fhd0 arbidopsis
157	17	68.0	540	2	0524H4_MAGGR	0524h4 magnaporthe	230	17	68.0	1072	2	04Q1A6_LEIMA	04q1a6 leishmania
158	17	68.0	540	2	07Y000_ORYSA	07y000 oryza sativ	231	17	68.0	1087	2	07Q2T6_GIALA	07q2t6 giardia lam
159	17	68.0	543	2	064391_SOLATU	064391 xanthomonas	232	17	68.0	1163	3	057XP4_9TRYP	057xp4 trypanosoma
160	17	68.0	560	2	08PFI1_XANAC	08pfi1 xanthomonas	233	17	68.0	1220	2	07SC14_NEUCR	07sc14 neurospora
161	17	68.0	561	2	09VJ24_DROME	09vj24 drosophila	234	17	68.0	1244	2	05R736_PONPY	05r736 pongo pygma
162	17	68.0	564	2	05YRZ7_NOCRA	05yrz7 nocardiella fa	235	17	68.0	1244	2	04P0V7_USTMA	04p0v7 ustilago ma
163	17	68.0	568	2	015968_DROME	015968 drosophila	236	17	68.0	1259	2	0570U5_9TRYP	0570u5 trypanosoma
164	17	68.0	568	2	09VEN8_DROME	09ven8 drosophila	237	17	68.0	1264	2	0570U52_9TRYP	0570u52 trypanosoma
165	17	68.0	568	2	08S9B4_VOLCA	08s9b4 volvox cart	238	17	68.0	1267	2	04H3A4_CIOIN	04h3a4 cloia inte
166	17	68.0	572	2	0691A3_ORYSA	0691a3 oryza sativ	239	17	68.0	1477	2	053MZ6_ORYSA	053mz6 oryza sativ
167	17	68.0	573	2	0948Y9_VOLCA	0948y9 volvox cart	240	17	68.0	1584	2	04T364_TERNG	04t364 tetraodon n
168	17	68.0	575	1	TREB_HUMAN	P07204 homo sapien	241	17	68.0	1625	2	ZCH11_HUMAN	05tcx3 homo sapien
169	17	68.0	575	1	TREB_SAISC	Q71U07 salmistr sci	242	17	68.0	1644	1	04PDV9_USTMA	04pdv9 ustilago ma
170	17	68.0	582	2	08IV29_HUMAN	08iv29 homo sapien	243	17	68.0	1737	2	05SU07_CRYNE	05su07 cryptococcu
171	17	68.0	582	2	06S153_CHLRE	06s153 chlamydomon	244	17	68.0	1858	2	08U0W7_CRYNE	08j0w7 cryptococcu
172	17	68.0	582	2	09SB11_CHLRE	09sb11 chlamydomon	245	17	68.0	1858	2	05KHN4_CRYNE	05khn4 cryptococcu
173	17	68.0	584	1	HMEN_DROVI	P091A5 drosophila	246	17	68.0	1858	2	04QC96_LEIMA	04qc96 leishmania
174	17	68.0	593	2	0818V8_GIALA	0818v8 giardia lam	247	17	68.0	2049	2	04QX96_LEIMA	04qx96 leishmania
175	17	68.0	593	2	TNR8_HUMAN	P28908 homo sapien	248	17	68.0	2103	2	05AUX1_EMENT	05aux1 aspergillus
176	17	68.0	596	2	086BD0_HALRO	086bd0 halocynthia	249	17	68.0	2259	2	0921C2_MOUSE	0921c2 mus musculu
177	17	68.0	600	2	09HZW1_PSEAE	Q9hzw1 pseudomonas	250	17	68.0	2304	2	0921C3_MOUSE	0921c3 mus musculu

251	17	68.0	2395	2	Q27167_PARTE	Q27167 paramecium	324	16	64.0	118	2	Q8RRE3_FUSNN	Q8RRE3 fusobacteri
252	17	68.0	2511	2	Q4NKK5_9DELT	Q4NKK5 anaeromyxob	325	16	64.0	120	2	Q5VZF6_HUMAN	Q5VZF6 homo sapien
253	17	68.0	2704	1	G168_PARRP	P17053 paramecium	326	16	64.0	120	2	Q7PDP29_GIALA	Q7PDP29 giardia lam
254	17	68.0	2715	1	G156_PARRP	P13837 paramecium	327	16	64.0	120	2	Q7P6A3_FUSNV	Q7P6A3 fusobacteri
255	17	68.0	2717	2	Q94710_PARTE	Q94710 paramecium	328	16	64.0	120	2	Q994S3_9PARA	Q994S3 snake atcc-
256	17	68.0	2729	2	Q6PQK6_PARTE	Q6PQK6 paramecium	329	16	64.0	120	2	Q4S916_TETNG	Q4S916 tetraodon n
257	17	68.0	2828	2	Q4STC7_TETNG	Q4STC7 tetraodon n	330	16	64.0	121	2	Q7R543_GIALA	Q7R543 giardia lam
258	17	68.0	3622	2	Q6BG85_PARTE	Q6BG85 paramecium	331	16	64.0	122	2	Q4THA6_TETNG	Q4THA6 tetraodon n
259	17	68.0	3982	2	Q4RV74_TETNG	Q4RV74 tetraodon n	332	16	64.0	123	1	CD59A_MOUSE	Q5S186 mus musculu
260	17	68.0	4588	2	Q4SOC2_TETNG	Q4SOC2 tetraodon n	333	16	64.0	123	1	Q5M6S2_TOBAC	Q5M6S2 nicotiana t
261	17	68.0	8243	2	Q96554_CRYEV	Q96554 crypospori	334	16	64.0	123	2	Q6ZDP91_ORYSA	Q6ZDP91 oryza sativ
262	17	68.0	8243	2	Q5COM8_CRYEV	Q5COM8 crypospori	335	16	64.0	123	2	Q6Q183_RAT	Q6Q183 rattus norv
263	16	64.0	38	2	Q5C643_SCHJA	Q5C643 schistosoma	336	16	64.0	123	2	Q920G6_MOUSE	Q920G6 mus musculu
264	16	64.0	42	2	Q4TDQ6_TETNG	Q4TDQ6 tetraodon n	337	16	64.0	123	2	Q542R7_MOUSE	Q542R7 mus musculu
265	16	64.0	51	2	Q5BTR4_SCHJA	Q5BTR4 schistosoma	338	16	64.0	123	2	Q4TIG5_TETNG	Q4TIG5 tetraodon n
266	16	64.0	51	2	Q56AS3_VYTRU	Q56AS3 bacillus th	339	16	64.0	124	2	Q8BSA5_MOUSE	Q8BSA5 mus musculu
267	16	64.0	52	2	Q9JWY4_NEITVA	Q9JWY4 neiseeria m	340	16	64.0	125	2	Q8RX30_RALSO	Q8RX30 ralestonia s
268	16	64.0	52	2	Q8JUN6_BRAJA	Q8JUN6 bradyrhizob	341	16	64.0	126	2	Q8WXA2_HUMAN	Q8WXA2 homo sapien
269	16	64.0	55	2	Q6LC75_RAT	Q6LC75 rattus norv	342	16	64.0	127	2	Q7SEF76_NEUCR	Q7SEF76 neuropept
270	16	64.0	56	2	Q7QY94_GIALA	Q7QY94 giardia lam	343	16	64.0	127	2	Q8TEB03_HUMAN	Q8TEB03 homo sapien
271	16	64.0	61	2	Q86YX3_HUMAN	Q86YX3 homo sapien	344	16	64.0	127	2	Q9H743_HUMAN	Q9H743 homo sapien
272	16	64.0	62	2	Q91FZ6_RRV6	Q91FZ6 chilo itide	345	16	64.0	128	1	LSHB_STRCA	Q6ZT54_HUMAN
273	16	64.0	63	2	Q8S5F8_ORYSA	Q8S5F8 oryza sativ	346	16	64.0	130	2	Q6VAH8_CANPA	Q6VAH8 canis famli
274	16	64.0	69	2	Q7R1B7_GIALA	Q7R1B7 giardia lam	347	16	64.0	130	2	Q5T565_ANOGA	Q5T565 anopheles g
275	16	64.0	71	2	Q4U002_XANCP	Q4U002 xanthomonas	348	16	64.0	131	2	Q67FZ6_ASF	Q67FZ6 aspergillu
276	16	64.0	72	2	Q6Z1G8_ORYSA	Q6Z1G8 oryza sativ	349	16	64.0	131	2	Q4PCR7_USTMA	Q4PCR7 ustilago ma
277	16	64.0	73	2	Q4TZZ0_9AVES	Q4TZZ0 anser anser	350	16	64.0	132	2	Q5RDPQ9_PONPY	Q5RDPQ9 pongo pygma
278	16	64.0	78	2	Q7UHI8_RHOBA	Q7UHI8 rhodopirell	351	16	64.0	132	2	Q5A116_CANAL	Q5A116 candida alb
279	16	64.0	80	2	Q50L66_DROBP	Q50L66 drosoephila	352	16	64.0	133	2	Q4S454_CAREL	Q4S454 caenorhabdi
280	16	64.0	80	2	Q50L69_DROBP	Q50L69 drosoephila	353	16	64.0	133	2	Q4B8V6_USTMA	Q4B8V6 ustilago ma
281	16	64.0	82	2	Q73JB4_TREDB	Q73JB4 treponema d	354	16	64.0	134	2	Q90W69_ONCMY	Q90W69 oncorhynch
282	16	64.0	87	2	Q9CGR4_LACIA	Q9CGR4 laccococcus	355	16	64.0	134	2	Q4PR87_USTMA	Q4PR87 ustilago ma
283	16	64.0	90	2	Q7ORS4_GIALA	Q7ORS4 giardia lam	356	16	64.0	135	2	Q6ATM6_ORYSA	Q6ATM6 oryza sativ
284	16	64.0	90	2	Q4RYA4_TETNG	Q4RYA4 tetraodon n	357	16	64.0	135	2	Q6H804_ORYSA	Q6H804 oryza sativ
285	16	64.0	91	1	UTER_HUMAN	Q6FHN3 homo sapien	358	16	64.0	136	2	Q5MCG5_9ANNE	Q5MCG5 mesenchytra
286	16	64.0	91	2	Q4XCA9_PLACH	Q4XCA9 plasmodium	359	16	64.0	136	2	Q6L6Q6_CRAIG	Q6L6Q6 crassostrea
287	16	64.0	92	2	Q8S9C2_BRAOL	Q8S9C2 brassica ol	360	16	64.0	137	2	Q9GSA5_CHOFU	Q9GSA5 choriostene
288	16	64.0	92	2	Q8RYV8_BRACM	Q8RYV8 brassica ca	361	16	64.0	138	2	Q5LCB4_BACFN	Q5LCB4 bacteroides
289	16	64.0	94	2	Q8T945_DROME	Q8T945 drosoephila	362	16	64.0	141	2	Q6ZDV3_HUMAN	Q6ZDV3 homo sapien
290	16	64.0	94	2	Q4S2A8_TETNG	Q4S2A8 tetraodon n	363	16	64.0	141	2	Q4TBM8_TETNG	Q4TBM8 tetraodon n
291	16	64.0	94	2	UTER_MOUSE	Q6FHN3 homo sapien	364	16	64.0	142	2	Q8BHR2_ORYSA	Q8BHR2 oryza sativ
292	16	64.0	96	1	UTER_RAT	Q6FHN3 homo sapien	365	16	64.0	142	2	Q8BDN0_VIBVU	Q8BDN0 vibrio vuln
293	16	64.0	96	1	UTER_RAT	Q6FHN3 homo sapien	366	16	64.0	142	2	Q7MNS9_VIBVU	Q7MNS9 vibrio vuln
294	16	64.0	96	2	Q7Y2V9_9CAUD	Q7Y2V9 stx2 conver	367	16	64.0	142	2	Q6DGP8_BRARE	Q6DGP8 brachydanio
295	16	64.0	96	2	Q7Y3C0_9CAUD	Q7Y3C0 stx1 conver	368	16	64.0	142	2	Q9ACU2_STRGO	Q9ACU2 streptomyce
296	16	64.0	96	2	Q8SCA2_9CAUD	Q8SCA2 stx2 conver	369	16	64.0	143	2	Q92WZ2_RHIME	Q92WZ2 rhizobium m
297	16	64.0	97	2	Q7RDA7_PLAYO	Q7RDA7 plasmodium	370	16	64.0	143	2	Q5XJ28_BRARE	Q5XJ28 brachydanio
298	16	64.0	98	2	Q5TAF6_HUMAN	Q5TAF6 homo sapien	371	16	64.0	144	2	Q5WLK3_BAC9K	Q5WLK3 bacillus cl
299	16	64.0	98	2	Q8ZD47_YERPE	Q8ZD47 yerinia pe	372	16	64.0	145	2	Q5UBD4_BERVU	Q5UBD4 beta vulgar
300	16	64.0	98	2	Q668U9_YERPS	Q668U9 yerinia ps	373	16	64.0	147	2	Q8B710_MOUSE	Q8B710 mus musculu
301	16	64.0	98	2	Q6D2I5_YERMCT	Q6D2I5 erwania car	374	16	64.0	147	2	Q9YAS3_AERPE	Q9YAS3 aeropyrum p
302	16	64.0	99	2	Q4T773_AZOVI	Q4T773 azotobacter	375	16	64.0	148	2	Q86V99_HUMAN	Q86V99 homo sapien
303	16	64.0	100	2	Q8D0U9_YERPE	Q8D0U9 yerinia pe	376	16	64.0	148	2	Q9SKD2_MACPA	Q9SKD2 macaca fasc
304	16	64.0	100	2	Q4KH59_PSEPS	Q4KH59 pseudomonas	377	16	64.0	148	2	Q9N018_MACPA	Q9N018 macaca fasc
305	16	64.0	101	2	Q6LPZ1_PHOBR	Q6LPZ1 photobacter	378	16	64.0	148	2	Q8PEU4_XANAC	Q8PEU4 xanthomonas
306	16	64.0	101	2	Q52XM6_HPBVO	Q52XM6 hepatitis b	379	16	64.0	148	2	Q9QSS5_9GBMT	Q9QSS5 sweet potat
307	16	64.0	103	2	P74449_SYNY3	P74449 synecocyst	380	16	64.0	149	2	Q64T88_BACFR	Q64T88 bacteroides
308	16	64.0	104	1	YR7H_ECOLI	YR7H escherichia	381	16	64.0	150	2	Q6E0M6_ORYSA	Q6E0M6 oryza sativ
309	16	64.0	107	2	Q67FZ2_ASF	Q67FZ2 african swi	382	16	64.0	151	2	Q4P3H1_USTMA	Q4P3H1 ustilago ma
310	16	64.0	107	2	Q67FZ3_ASF	Q67FZ3 african swi	383	16	64.0	151	2	Q5FC01_GLUOX	Q5FC01 gluconobact
311	16	64.0	107	2	Q67FZ4_ASF	Q67FZ4 african swi	384	16	64.0	152	2	Q6VTN9_ORYSA	Q6VTN9 oryza sativ
312	16	64.0	109	2	Q6XAU4_9BACT	Q6XAU4 uncultured	385	16	64.0	152	2	Q8USM5_AGRTS	Q8USM5 agrobacteri
313	16	64.0	111	2	Q4MOE3_9BUPK	Q4MOE3 burkholderi	386	16	64.0	153	1	BD8C_DEIRA	Q9RUB5 delnococtus
314	16	64.0	111	2	Q4CTC3_TETNG	Q4CTC3 tetraodon n	387	16	64.0	153	1	Q67858_HPBVO	Q67858 hepatitis b
315	16	64.0	112	1	Y1386_PASWU	Y1386 pasteurella	388	16	64.0	154	2	Q63TR9_BURPS	Q63TR9 burkholderi
316	16	64.0	112	2	Q7R5D2_GIALA	Q7R5D2 giardia lam	389	16	64.0	155	2	O16422_CAREL	O16422 caenorhabdi
317	16	64.0	112	2	Q4TOS8_9SPFN	Q4TOS8 erythrobact	390	16	64.0	156	2	O5YVL9_NOCPA	O5YVL9 nocardiia fa
318	16	64.0	115	2	Q5MXR8_LEBPL	Q5MXR8 legionella	391	16	64.0	157	2	Q9BH65_GIALA	Q9BH65 giardia lam
319	16	64.0	115	2	Q5X6A6_LEGPA	Q5X6A6 legionella	392	16	64.0	157	2	Q9BIL6_GIALA	Q9BIL6 giardia lam
320	16	64.0	116	1	Q7V6U7_PROMM	Q7V6U7 prochloroco	393	16	64.0	157	2	Q6D092_ORYSA	Q6D092 oryza sativ
321	16	64.0	116	1	UL81_HCIVA	UL81 human cytom	394	16	64.0	157	2	Q7X094_9BACT	Q7X094 uncultured
322	16	64.0	118	2	Q8TUD7_METAC	Q8TUD7 methanosarc	395	16	64.0	158	2	Q9BIL8_GIALA	Q9BIL8 giardia lam
323	16	64.0	118	2	Q525Q2_MAGGR	Q525Q2 magnaporthe	396	16	64.0	158	2		

397	16	64.0	158	2	07XD88_ORYSA	07XD88_oryza sativ	470	16	64.0	184	2	04J207_AZOVI	04J207_azotobacter
398	16	64.0	158	2	08W375_ORYSA	08W375_oryza sativ	471	16	64.0	184	2	04KJ04_PSEFS	04KJ04_pseudomonas
399	16	64.0	158	2	07LZD1_CHICK	07LZD1_gallus gall	472	16	64.0	184	2	06AK89_PROAC	06AK89_propionibac
400	16	64.0	159	1	LSHB_MELGA	P45646 melaeagris g	473	16	64.0	186	2	06B865_9ACAR	06B865_ixodes paci
401	16	64.0	159	2	05P149_AZOSU	05P149 azoarcus sp	474	16	64.0	186	2	04RR86_TETNG	04RR86_tetradon n
402	16	64.0	161	2	05KSB7_OLIMA	05KSB7 oligobatrach	475	16	64.0	186	2	06BK05_9SAUR	06BK05_chrysomya p
403	16	64.0	161	2	04GY18_9TRYEP	04GY18 trypanosoma	476	16	64.0	187	2	04MOE2_9BURK	04MOE2_burkholderi
404	16	64.0	161	2	09KKJ0_YEREN	09KKJ0 yerenia en	477	16	64.0	189	2	04T3W8_TETNG	04T3W8_tetradon n
405	16	64.0	163	2	0967R9_GIALA	0967R9 giardia lam	478	16	64.0	190	2	05AP86_CANAL	05AP86_candida alb
406	16	64.0	164	2	06J3K8_BURFS	06J3K8 burkholderi	479	16	64.0	190	2	007525_TROMA	007525_tropaeolum
407	16	64.0	164	2	07NV44_CHRVO	07NV44 chroomobacte	480	16	64.0	190	2	06R6N2_BRUTL	06R6N2_eucomia ul
408	16	64.0	164	2	062DA8_BURMA	062DA8 burkholderi	481	16	64.0	190	2	06GNTC_BRATA	06GNTC_bredytriazob
409	16	64.0	165	2	07PCMO_GIALA	07PCMO giardia lam	482	16	64.0	191	2	07QB93_ANOGA	07QB93_anopheles g
410	16	64.0	165	2	06K926_ORYSA	06K926 oryza sativ	483	16	64.0	192	2	05DTE9_CABEL	05DTE9_caenorhabdi
411	16	64.0	166	1	LSHB_COTUA	P45657 coturnix co	484	16	64.0	192	2	08H7M5_ORYSA	08H7M5_oryza sativ
412	16	64.0	166	1	004434_YEAST	004434 saccharomyc	485	16	64.0	195	2	098KC2_RHITO	098KC2_rhizobium l
413	16	64.0	166	2	001762_CABEL	001762 caenorhabdi	486	16	64.0	198	2	04L170_9BURK	04L170_burkholderi
414	16	64.0	166	2	09FW36_ORYSA	09FW36 oryza sativ	487	16	64.0	199	2	07QL69_ANOGA	07QL69_anopheles g
415	16	64.0	166	2	08BDC6_HCMV	08BDC6 human cytom	488	16	64.0	200	2	04LJ73_9BURK	04LJ73_burkholderi
416	16	64.0	167	2	06ZRS3_HUMAN	06ZRS3 homo sapien	489	16	64.0	200	2	06MJ55_BDEBA	06MJ55_bdeilovibri
417	16	64.0	167	2	06ZU83_HUMAN	06ZU83 homo sapien	490	16	64.0	201	2	04TAP4_GIBZE	04TAP4_gibberella
418	16	64.0	167	2	04RUS7_TETNG	04RUS7 tetradon n	491	16	64.0	201	2	05SQ62_HUMAN	05SQ62_homo sapien
419	16	64.0	168	2	05RBM5_PONPY	05RBM5 pongo pygma	492	16	64.0	201	2	08BDX3_HUMAN	08BDX3_homo sapien
420	16	64.0	168	2	08BRM9_MOUSE	08BRM9 mus musculu	493	16	64.0	202	2	05VZF9_HUMAN	05VZF9_homo sapien
421	16	64.0	169	2	059W12_CANAL	059W12 candida alb	494	16	64.0	202	2	05VNM4_ORYSA	05VNM4_oryza sativ
422	16	64.0	171	1	TRAY_ECOLI	P41069 escherichia	495	16	64.0	203	2	08KNL2_SALTI	08KNL2_salmonella
423	16	64.0	171	1	05UBM7_ECOLI	05UBM7 escherichia	496	16	64.0	204	2	05P8N7_AZOSE	05P8N7_azoarcus sp
424	16	64.0	171	2	032532_ECOLI	032532 escherichia	497	16	64.0	204	2	07OS93_GIALA	07OS93_giardia lam
425	16	64.0	171	2	084A31_ECOLI	084A31 escherichia	498	16	64.0	206	2	04SKY8_TETNG	04SKY8_tetradon n
426	16	64.0	171	2	07AK72_92ZIZE	07AK72 plasmid r10	499	16	64.0	206	2	028702_ARCFU	028702_arthaeoglob
427	16	64.0	172	2	04HKX5_GIBZE	04HKX5 gibberella	500	16	64.0	207	2	04VDG6_9STRPA	04VDG6_hyaloperono
428	16	64.0	172	2	05HX08_GLUOX	05HX08 gluconobact	501	16	64.0	207	2	05CY78_CRYPP	05CY78_cryptospori
429	16	64.0	173	2	04XQ15_PLACH	P22389 mus musculu	502	16	64.0	208	2	05CL06_CRYHO	05CL06_cryptospori
430	16	64.0	173	1	EDN2_MOUSE	0910h9 human cytom	503	16	64.0	208	2	07OSQ7_GIALA	07OSQ7_giardia lam
431	16	64.0	175	2	0910H9_HCMV	0910H9 human cytom	504	16	64.0	209	2	05OSQ7_ENTH1	05OSQ7_entamoeba h
432	16	64.0	175	2	091IR1_HCMV	091IR1 human cytom	505	16	64.0	209	2	05NT08_ENTH1	05NT08_entamoeba h
433	16	64.0	175	2	05RLB0_HCMV	05RLB0 human cytom	506	16	64.0	209	2	051263_MACGR	051263_magnetoporth
434	16	64.0	175	2	05RLB3_HCMV	05RLB3 human cytom	507	16	64.0	209	2	04P079_USTMA	04P079_xenopus lae
435	16	64.0	175	2	05RLB4_HCMV	05RLB4 human cytom	508	16	64.0	209	2	070050_XENIA	070050_xenopus lae
436	16	64.0	175	2	05RLB5_HCMV	05RLB5 human cytom	509	16	64.0	210	2	06UXP6_HUMAN	06UXP6_homo sapien
437	16	64.0	175	2	05RLD5_HCMV	05RLD5 human cytom	510	16	64.0	210	2	08FXQ3_BRUSU	08FXQ3_brucella ab
438	16	64.0	175	2	09PWY0_HCMV	09PWY0 human cytom	511	16	64.0	210	2	04SIVA_TETNG	04SIVA_tetradon n
439	16	64.0	175	2	09PXA1_HCMV	09PXA1 human cytom	512	16	64.0	211	2	051263_MACGR	051263_magnetoporth
440	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	513	16	64.0	212	2	05SRO2_HUMAN	05SRO2_homo sapien
441	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	514	16	64.0	212	2	05SRO2_HUMAN	05SRO2_homo sapien
442	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	515	16	64.0	212	2	05SST4_HUMAN	05SST4_homo sapien
443	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	516	16	64.0	212	2	07ONF1_ANOGA	07ONF1_anopheles g
444	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	517	16	64.0	212	2	07SCN6_NEUCR	07SCN6_neutrospora
445	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	518	16	64.0	214	2	04P079_USTMA	04P079_xenopus lae
446	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	519	16	64.0	214	2	070050_XENIA	070050_xenopus lae
447	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	520	16	64.0	214	2	06UXP6_HUMAN	06UXP6_homo sapien
448	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	521	16	64.0	218	2	08R2V6_MOUSE	08R2V6_mus musculu
449	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	522	16	64.0	218	2	08R2V6_MOUSE	08R2V6_mus musculu
450	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	523	16	64.0	219	2	08R2V6_MOUSE	08R2V6_mus musculu
451	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	524	16	64.0	219	2	08R2V6_MOUSE	08R2V6_mus musculu
452	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	525	16	64.0	220	2	05GAL1_9YIRU	05GAL1_griouper iti
453	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	526	16	64.0	220	2	0580C5_9TRYEP	0580C5_entramoeba h
454	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	527	16	64.0	222	2	050219_ENTH1	050219_entramoeba h
455	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	528	16	64.0	222	2	04KIM9_PSEFS	04KIM9_pseudomonas
456	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	529	16	64.0	223	2	07QX05_GIALA	07QX05_giardia lam
457	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	530	16	64.0	224	2	09CS01_MOUSE	09CS01_mus musculu
458	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	531	16	64.0	225	2	08CD47_MOUSE	08CD47_mus musculu
459	16	64.0	175	2	09PZQ6_HCMV	09PZQ6 human cytom	532	16	64.0	226	2	09UVCS_SYNRA	09UVCS_syncepalas
460	16	64.0	176	1	EDN2_RAT	P23943 rattus norv	533	16	64.0	230	1	Y4IC_RHISN	Y54486_rhizobium s
461	16	64.0	176	2	08CEX7_MOUSE	08CEX7 mus musculu	534	16	64.0	230	2	06S9V4_ORYSA	06S9V4_oryza sativ
462	16	64.0	177	2	08XZC1_RALSO	08XZC1 ralestonia s	535	16	64.0	232	2	07SGR1_ORYSA	07SGR1_oryza sativ
463	16	64.0	178	2	06ZW12_HUMAN	06ZW12 homo sapien	536	16	64.0	232	2	04RE98_TETNG	04RE98_tetradon n
464	16	64.0	179	2	06SAH6_YERPE	06SAH6 yerenia pe	537	16	64.0	233	2	0946Y8_HORVU	0946Y8_hordium vul
465	16	64.0	180	2	06YW25_ORYSA	06YW25 oryza sativ	538	16	64.0	235	1	NKX28_MOUSE	070584_mus musculu
466	16	64.0	182	2	082494_ARATH	082494 arabidopsis	539	16	64.0	235	2	04LEI7_BRYMA	04LEI7_bryopsis ma
467	16	64.0	183	2	053EX3_HUMAN	053EX3 homo sapien	540	16	64.0	236	2	08BMW7_MOUSE	08BMW7_m mus muscu
468	16	64.0	183	2	04J2P2_AZOVI	04J2P2 azotobacter	541	16	64.0	237	2	04IW40_GIBZE	04IW40_gibberella
469	16	64.0	184	2	086PE0_DROME	086PE0 drosophila	542	16	64.0	238	2	070U60_GIALA	07QU60_giardia lam

543	16	64.0	238	2	004364_ORYSA	004364_oryza sativ	616	16	64.0	284	2	096MVO_HUMAN	096MVO_homo sapien
544	16	64.0	238	2	067207_LYCES	067207_lycopersico	617	16	64.0	285	2	07R3E3_GIALA	07R3E3_giardia lam
545	16	64.0	239	2	09NK17_HUMAN	09NK17_homo sapien	618	16	64.0	285	2	04P024_9DELT	04P024_anaeromyxob
546	16	64.0	239	2	07R376_GIALA	07R376_giardia lam	619	16	64.0	285	2	09DGH7_CHICK	09DGH7_gallus gall
547	16	64.0	240	2	04P847_USTMA	04P847_ustillago ma	620	16	64.0	286	2	07D0X6_AGR75	07D0X6_agrobacteri
548	16	64.0	240	2	063MM3_BURPS	063MM3_burholderi	621	16	64.0	286	2	09DBD2_M009E	09DBD2_mus musculu
549	16	64.0	242	2	051XD3_MAGOR	051XD3_magaporthe	622	16	64.0	288	2	0526H0_ORYJA	0526H0_oryza sativ
550	16	64.0	243	2	08BJC3_MOUSE	08BJC3_mus musculu	623	16	64.0	289	2	09RXY6_DEIRA	09RXY6_dainococcus
551	16	64.0	243	2	083290_ORYLA	083290_oryza lat	624	16	64.0	291	2	07Q3X6_ANGDA	07Q3X6_anopheles g
552	16	64.0	244	2	05B708_EMENI	05B708_aspergillus	625	16	64.0	291	2	04Z039_PLAAB	04Z039_plasmodium
553	16	64.0	244	2	095ZL6_CAEBL	095ZL6_caenorhabdi	626	16	64.0	291	2	05VNP0_ORYSA	05VNP0_oryza sativ
554	16	64.0	246	2	062126_CAEBL	062126_caenorhabdi	627	16	64.0	295	2	007524_ROMA	007524_tropaeolum
555	16	64.0	246	2	04Z0Q4_PLABE	04Z0Q4_plasmodium	628	16	64.0	295	2	053NC6_ORYSA	053NC6_oryza sativ
556	16	64.0	246	2	06LND8_PHOPR	06LND8_phocobacter	629	16	64.0	296	2	05OLC2_9DIPT	05OLC2_drosophila
557	16	64.0	246	2	08YDU4_BRUME	08YDU4_brucella me	630	16	64.0	296	2	05OLB1_DROBP	05OLB1_drosophila
558	16	64.0	247	2	096CB6_CAEBL	096CB6_caenorhabdi	631	16	64.0	296	2	05OLM0_DROBP	05OLM0_drosophila
559	16	64.0	247	2	04UT88_XANCP	04UT88_xanthomonas	632	16	64.0	296	2	05OLA5_9DIPT	05OLA5_drosophila
560	16	64.0	247	2	08PAD7_XANCP	08PAD7_xanthomonas	633	16	64.0	296	2	05OLA8_9DIPT	05OLA8_drosophila
561	16	64.0	247	2	092326_MOUSE	092326_mus musculu	634	16	64.0	296	2	05OLA7_9DIPT	05OLA7_drosophila
562	16	64.0	248	2	061HW9_CAEBR	061HW9_caenorhabdi	635	16	64.0	296	2	05OLA6_9DIPT	05OLA6_drosophila
563	16	64.0	248	2	045450_CAEBL	045450_caenorhabdi	636	16	64.0	296	2	05OLA5_9DIPT	05OLA5_drosophila
564	16	64.0	248	2	09D1R7_MOUSE	09D1R7_mus musculu	637	16	64.0	296	2	05OLA4_9DIPT	05OLA4_drosophila
565	16	64.0	249	2	04I4U5_GIBZE	04I4U5_gibberella	638	16	64.0	296	2	05OLA3_9DIPT	05OLA3_drosophila
566	16	64.0	249	2	05YPM1_NOCFA	05YPM1_nocardia fa	639	16	64.0	297	2	05OLA2_9DIPT	05OLA2_drosophila
567	16	64.0	249	2	091686_XENLA	091686_xenopus lae	640	16	64.0	297	2	05VYX0_HUMAN	05VYX0_homo sapien
568	16	64.0	252	2	095ZL7_CAEBL	095ZL7_caenorhabdi	641	16	64.0	297	2	04VSR0_DROME	04VSR0_drosophila
569	16	64.0	252	2	09SZU5_ARATH	09SZU5_arabidopsis	642	16	64.0	298	2	055HN3_CRYNE	055HN3_cryptococcu
570	16	64.0	252	2	08BX9_SHRON	08BX9_shewanella	643	16	64.0	298	2	05K773_CRYNE	05K773_cryptococcu
571	16	64.0	253	2	07R96_BOVIN	07R96_bos taurus	644	16	64.0	299	1	XTB32_ARATH	09B119_arabidopsis
572	16	64.0	253	2	09FY31_FESPR	09FY31_festuca pra	645	16	64.0	299	1	04V5X8_DROME	04V5X8_drosophila
573	16	64.0	254	2	05LUB1_SILPO	05LUB1_silicibacte	646	16	64.0	299	2	08GZFL_LYCES	08GZFL_lycopersico
574	16	64.0	254	2	05N1Y1_SYMP6	05N1Y1_synchococc	647	16	64.0	299	2	09HVX9_PSEAE	09HVX9_pseudomonas
575	16	64.0	255	1	TNR9_HUMAN	007011_homo sapien	648	16	64.0	300	2	060960_LEIMA	060960_leishmania
576	16	64.0	256	2	04Q6F4_HAB18	04Q6F4_haemophilus	649	16	64.0	301	2	06CX25_KUTIA	06CX25_kluyveromyc
577	16	64.0	256	2	06ON66_CAEBR	06ON66_caenorhabdi	650	16	64.0	301	2	06Z4N0_ORYJA	06Z4N0_oryza sativ
578	16	64.0	258	2	006BG3_PARTE	006BG3_parametium	651	16	64.0	302	2	05CND4_CRYHO	05CND4_cryptospori
579	16	64.0	258	2	000946_9HYMN	000946_tetrahymena	652	16	64.0	302	2	07YYG0_CRYPV	07YYG0_cryptospori
580	16	64.0	258	2	082710_GALPH	082710_gallid herp	653	16	64.0	303	2	06W743_CRYPV	06W743_cryptospori
581	16	64.0	259	2	051PM9_MAGGR	051PM9_magnaporthe	654	16	64.0	303	2	06W744_CRYPV	06W744_cryptospori
582	16	64.0	261	1	GSHI_MOUSE	P31315_mus musculu	655	16	64.0	303	2	04Q4P9_LEIMA	04Q4P9_leishmania
583	16	64.0	261	1	Y118_HABIN	057097_haemophilus	656	16	64.0	306	2	0821D8_CHICV	0821D8_chlamydomph
584	16	64.0	262	1	VMEL_CVPFS	P09175_porcine tira	657	16	64.0	306	2	08N999_PSEBP	08N999_pseudomonas
585	16	64.0	264	1	GSHI_HUMAN	09H482_homo sapien	658	16	64.0	307	2	04WXG7_ASFPFL	04WXG7_aspergillus
586	16	64.0	264	1	067WK3_ORYSA	067WK3_oryza sativ	659	16	64.0	309	2	09N2P0_9CILI	09N2P0_parametium
587	16	64.0	265	2	097EL5_CIOAB	097EL5_clostridium	660	16	64.0	307	2	06CTM4_HUMAN	06CTM4_homo sapien
588	16	64.0	267	2	090257_FLDV	090257_fish lympho	661	16	64.0	308	2	07R414_GIALA	07R414_giardia lam
589	16	64.0	268	1	HISU_NEIGO	006758_neisseria g	662	16	64.0	308	2	06C2H3_YARLI	06C2H3_yarrowia il
590	16	64.0	268	2	097AD5_THEVO	097AD5_thermoplaem	663	16	64.0	309	2	05O1U0_ORYJA	05O1U0_oryza sativ
591	16	64.0	268	2	05F718_NEIG1	05F718_neisseria g	664	16	64.0	309	2	091T37_LSDV	091T37_lumpy skin
592	16	64.0	268	2	09UTB4_NEIMA	09UTB4_neisseria m	665	16	64.0	310	1	TH14_HALSA	091unc7_halobacteri
593	16	64.0	268	2	09JYF0_NEIMB	09JYF0_neisseria m	666	16	64.0	310	2	0813O5_PLA7	0813O5_plasmodium
594	16	64.0	269	1	THID_STRGO	092066_streptomyces	667	16	64.0	310	2	09E2Q7_NOCAS	09E2Q7_nocardia as
595	16	64.0	270	2	08MUT0_TIGCA	08MUT0_ligilopus c	668	16	64.0	310	2	07V249_PROMP	07V249_prochloroco
596	16	64.0	272	2	04F8P7_FUGRU	04F8P7_fugu rubrip	669	16	64.0	310	2	08PBR3_XANVC	08PBR3_xanthomonas
597	16	64.0	273	2	054QZ0_DICDI	054QZ0_diciclostei	670	16	64.0	310	2	04WOC8_ASFPFL	04WOC8_aspergillus
598	16	64.0	273	2	04TQTR_9SPHN	04TQTR_erythrobact	671	16	64.0	311	2	05H6C2_XANVR	05H6C2_xanthomonas
599	16	64.0	273	2	05RRU5_BRARE	05RRU5_baccharianu	672	16	64.0	312	2	072AR2_DESYR	072AR2_dentulfovibr
600	16	64.0	275	2	044824_CAEBL	044824_caenorhabdi	673	16	64.0	312	2	09DGB8_CHICK	09DGB8_gallus gall
601	16	64.0	275	2	09CS10_MOUSE	09CS10_mus musculu	674	16	64.0	312	2	06W955_NEUCR	06W955_neurospora
602	16	64.0	276	1	FXLZ20_RAT	09G347_rattus norv	675	16	64.0	313	2	07ONT6_GIALA	07ONT6_giardia lam
603	16	64.0	276	2	065235_ARATH	065235_arabidopsis	676	16	64.0	315	2	07TTV7_SYVPX	07TTV7_synchococc
604	16	64.0	277	2	063Y45_BURPS	063Y45_burholderi	677	16	64.0	315	2	07NM80_GLOVI	07NM80_gloeobacter
605	16	64.0	277	2	08KD78_CHL7E	08KD78_chlorobium	678	16	64.0	315	2	051875_HUMAN	051875_homo sapien
606	16	64.0	277	2	006477_MOUSE	006477_mus musculu	679	16	64.0	317	2	05VYX3_HUMAN	05VYX3_homo sapien
607	16	64.0	278	1	YOX1_SCHPO	060098_schizosacch	680	16	64.0	318	2	07NTB8_CARVO	07NTB8_chromobacte
608	16	64.0	278	2	07SCA7_NEUCR	07SCA7_neurospora	681	16	64.0	319	2	05CUP4_CRYPV	05CUP4_cryptospori
609	16	64.0	279	2	051IF7_MAGGR	051IF7_magnaporthe	682	16	64.0	319	2	08WST3_CIOIN	08WST3_ciona intes
610	16	64.0	280	2	05B665_BMENI	05B665_aspergillus	683	16	64.0	320	2	05CHH6_CRYHO	05CHH6_cryptospori
611	16	64.0	280	2	073BET3_MYCPA	073BET3_mycobacteri	684	16	64.0	320	2	07WJ16_BORPE	07WJ16_bordetella
612	16	64.0	280	2	073BET3_MYCPA	073BET3_mycobacteri	685	16	64.0	320	2	07WJ16_BORPE	07WJ16_bordetella
613	16	64.0	281	2	06LSD9_ORYSA	06LSD9_oryza sativ	686	16	64.0	320	2	059EP9_HUMAN	059EP9_homo sapien
614	16	64.0	283	2	074Z18_ASHGO	074Z18_ashya goss	687	16	64.0	321	2	061EH2_CAEBR	061EH2_caenorhabdi
615	16	64.0	283	2	09XSZ8_CERAE	09XSZ8_cercopithec	688	16	64.0	321	2		

689	16	64.0	321	2	Q7W9Z1 BORPA	Q7W9Z1 bordetella	762	16	64.0	379	2	Q043Z6_PICAN	Q043Z6 pichia angu
690	16	64.0	323	2	Q5UIQ1_ORYSA	Q5UIQ1 oryza sativ	763	16	64.0	381	2	Q59XB1_CANAL	Q59XB1 candida alb
691	16	64.0	323	2	Q9FRE1_ORYSA	Q9FRE1 oryza sativ	764	16	64.0	382	1	FERUB_HUMAN	Q99gms homo sapien
692	16	64.0	323	2	Q9FRE3_ORYSA	Q9FRE3 oryza sativ	765	16	64.0	382	2	Q59XF7_CANAL	Q59XF7 candida alb
693	16	64.0	325	1	Q8BUI6_MOUSE	Q8BUI6 mus musculu	766	16	64.0	382	2	Q6DKS6_HUMAN	Q6DKS6 homo sapien
694	16	64.0	327	1	TNR6_MOUSE	P25346 mus musculu	767	16	64.0	382	2	Q6GRB6_HUMAN	Q6GRB6 homo sapien
695	16	64.0	327	2	Q89FPH_BRAJA	Q89FPH bradyrhizob	768	16	64.0	382	2	Q6ZBF3_ORYSA	Q6ZBF3 oryza sativ
696	16	64.0	327	2	Q6GRF1_MOUSE	Q6GRF1 mus musculu	769	16	64.0	382	2	Q4RJ19_TERNG	Q4RJ19 tetradon n
697	16	64.0	329	2	Q7SBD9_NEUCR	Q7SBD9 neurospora	770	16	64.0	383	1	BSC12_MOUSE	Q92zef mus musculu
698	16	64.0	331	2	Q83359_TREPA	Q83359 treponema p	771	16	64.0	384	2	Q81208_PLAF7	Q81208 plasmodium
699	16	64.0	332	2	Q4KCG9_PSEFS	Q4KCG9 pseudomonas	772	16	64.0	386	2	Q4SCA3_TERNG	Q4SCA3 tetradon n
700	16	64.0	333	1	PEX2_ARATH	Q9C486 arabidopsis	773	16	64.0	387	2	Q7YUB7_BORPB	Q7YUB7 bordetella
701	16	64.0	333	2	Q73560_MYCPA	Q73560 mycobacteri	774	16	64.0	387	2	Q7WPC9_BORBR	Q7WPC9 borrelia
702	16	64.0	334	2	Q84UD6_ARATH	Q84UD6 arabidopsis	775	16	64.0	388	2	Q84RH8_CHLIS6	Q84RH8 chlorarachn
703	16	64.0	334	2	Q4XND1_PSESY	Q4XND1 pseudomonas	776	16	64.0	388	2	Q631R1_BURPS	Q631R1 burkholderi
704	16	64.0	335	2	Q615B5_ORYSA	Q615B5 oryza sativ	777	16	64.0	389	2	Q95018_CAEEL	Q95018 caenorhabdi
705	16	64.0	336	2	Q5H6C0_XANOR	Q5H6C0 xanthomonas	778	16	64.0	392	2	Q31277_RAT	Q31277 rattus norv
706	16	64.0	336	2	Q6LRF9_PHOPR	Q6LRF9 photobacter	779	16	64.0	394	2	Q4ICX6_GIBZE	Q4ICX6 gibberella
707	16	64.0	336	2	Q55908_SYNY3	Q55908 synecocyst	780	16	64.0	394	2	Q54PE7_DICDI	Q54PE7 dictyostei
708	16	64.0	336	2	Q87WM6_PSEBM	Q87WM6 pseudomonas	781	16	64.0	394	2	Q8BVH4_MOUSE	Q8BVH4 mus musculu
709	16	64.0	337	2	Q5N412_SYNP6	Q5N412 synecococc	782	16	64.0	395	2	Q800K8_PAROL	Q800K8 paratrichth
710	16	64.0	338	2	Q4Q7J8_LEIMA	Q4Q7J8 leishmania	783	16	64.0	396	2	Q4FYJ0_LEIMA	Q4FYJ0 leishmania
711	16	64.0	338	2	Q67TOS_SYNTM	Q67TOS symbiobacte	784	16	64.0	398	2	Q53BN3_HUMAN	Q53BN3 homo sapien
712	16	64.0	339	2	Q7PZW5_ANOGA	Q7PZW5 anopheles g	785	16	64.0	398	2	Q5CTJ2_HUMAN	Q5CTJ2 homo sapien
713	16	64.0	339	2	Q6NMQ7_BDEBA	Q6NMQ7 bdellovibri	786	16	64.0	400	1	BSC12_HUMAN	Q567S1_HUMAN
714	16	64.0	340	2	Q61M46_PHOPR	Q61M46 photobacter	787	16	64.0	400	2	Q6LWR7_METMP	Q6LWR7 methanococ
715	16	64.0	341	1	PTAFR_MOUSE	Q62035 mus musculu	788	16	64.0	401	2	Q7G7C9_ORYSA	Q7G7C9 oryza sativ
716	16	64.0	341	1	PTAFR_RAT	P46002 rattus norv	789	16	64.0	401	2	Q80WK4_MOUSE	Q80WK4 mus musculu
717	16	64.0	341	2	Q4UTJ4_CORAK	Q4UTJ4 corynebacte	790	16	64.0	402	2	Q5HDM2_XANOR	Q5HDM2 xanthomonas
718	16	64.0	341	2	Q8UTX8_LSDV	Q8UTX8 lumpy skin	791	16	64.0	402	2	Q8SBS6_DROVI	Q8SBS6 drophi
719	16	64.0	343	2	Q40148_LYCSB	Q40148 lycopersici	792	16	64.0	404	2	Q7WZ26_LEPIC	Q7WZ26 lepic
720	16	64.0	344	2	Q7QZRS_GIALA	Q7QZRS giardia lam	793	16	64.0	404	2	Q8EP94_LEBIN	Q8EP94 leishmania
721	16	64.0	344	2	Q8ZLT7_STRAM	Q8ZLT7 streptomyce	794	16	64.0	404	2	Q4V91L_BRARE	Q4V91L brachydanio
722	16	64.0	345	2	Q5J876_HUMAN	Q5J876 homo sapien	795	16	64.0	405	2	Q6NMK7_BDEBA	Q6NMK7 bdellovibri
723	16	64.0	345	2	Q73TU0_MYCPA	Q73TU0 mycobacteri	796	16	64.0	408	2	Q8BFW1_MOUSE	Q8BFW1 mus muscu
724	16	64.0	346	2	Q6Z4S3_ORYSA	Q6Z4S3 oryza sativ	797	16	64.0	408	2	Q8BMTT_MOUSE	Q8BMTT mus musculu
725	16	64.0	347	2	Q8SYW5_DROME	Q8SYW5 drosophila	798	16	64.0	408	2	Q8NK84_PHONA	Q8NK84 photiocta na
726	16	64.0	347	2	Q5OD15_MOUSE	Q5OD15 mus musculu	799	16	64.0	409	2	Q4NSK4_CAEEL	Q4NSK4 caenorhabdi
727	16	64.0	348	2	Q9RYI8_DEIRA	Q9RYI8 deinococcus	800	16	64.0	411	2	Q5YUD1_NOCPA	Q5YUD1 nocardia fa
728	16	64.0	348	2	Q8ZSB4_STRAV	Q8ZSB4 streptomyce	801	16	64.0	413	2	Q7OH00_ANOGA	Q7OH00 anopheles g
729	16	64.0	350	2	Q7N6Z3_GLOVI	Q7N6Z3 gloebacter	802	16	64.0	414	2	TNR16_CHICK	TNR16 gallus gall
730	16	64.0	351	2	Q81QV7_DROME	Q81QV7 drosophila	803	16	64.0	416	1	Q6ZBK5_CABBR	Q6ZBK5 caenorhabdi
731	16	64.0	351	2	Q8MR04_DROME	Q8MR04 drosophila	804	16	64.0	417	2	Q5VZF7_HUMAN	Q5VZF7 homo sapien
732	16	64.0	351	2	Q4RN41_TETNG	Q4RN41 tetradon n	805	16	64.0	418	2	Q9H599_HUMAN	Q9H599 homo sapien
733	16	64.0	351	2	Q4TDL0_TETNG	Q4TDL0 tetradon n	806	16	64.0	418	2	Q21941_CAEEL	Q21941 caenorhabdi
734	16	64.0	355	1	RHOW_DROME	P20350 drosophila	807	16	64.0	418	2	Q9FPA0_ORYSA	Q9FPA0 oryza sativ
735	16	64.0	355	2	Q540V7_DROME	Q540V7 drosophila	808	16	64.0	418	2	Q89SV5_BRAJA	Q89SV5 bradyrhizob
736	16	64.0	355	2	Q9AXJ5_CAEEL	Q9AXJ5 caenorhabdi	809	16	64.0	419	2	Q4RQ11_TETNG	Q4RQ11 tetradon n
737	16	64.0	355	2	Q94LQ4_ORYSA	Q94LQ4 oryza sativ	810	16	64.0	422	2	Q5LAS0_BACFR	Q5LAS0 bacteroides
738	16	64.0	356	2	Q8H8C7_ORYSA	Q8H8C7 oryza sativ	811	16	64.0	422	2	Q64QG8_BACFR	Q64QG8 bacteroides
739	16	64.0	357	2	Q93SK6_MYXXA	Q93SK6 myxococcus	812	16	64.0	422	1	TNR19_HUMAN	TNR19 homo sapien
740	16	64.0	357	2	Q87685_SIVCZ	Q87685 chimpanzee	813	16	64.0	423	1	Q5VZFB_HUMAN	Q5VZFB homo sapien
741	16	64.0	351	2	Q7OKX5_ANOGA	Q7OKX5 anopheles g	814	16	64.0	423	2	Q4IJUS_9HURK	Q4IJUS burkholderi
742	16	64.0	352	2	Q5TWJ4_MACMO	Q5TWJ4 macaca mula	815	16	64.0	423	2	Q9VU22_DROME	Q9VU22 drosophila
743	16	64.0	352	2	Q5BOY5_VIBEL	Q5BOY5 vibrio fisc	816	16	64.0	425	2	Q7VZ43_PROME	Q7VZ43 prochloroco
744	16	64.0	362	2	Q9X852_STRCO	Q9X852 streptomyce	817	16	64.0	425	2	Q5DIT5_MOUSE	Q5DIT5 mus musculu
745	16	64.0	363	2	Q7S9Y7_NEUCR	Q7S9Y7 neurospora	818	16	64.0	426	2	Q7RX63_NEUCR	Q7RX63 neurospora
746	16	64.0	363	2	Q9VW94_DROME	Q9VW94 drosophila	819	16	64.0	427	2	Q5CRG0_CRYPV	Q5CRG0 cryptospori
747	16	64.0	363	2	Q9BJ44_LEIMA	Q9BJ44 leishmania	820	16	64.0	428	2	Q5CGM9_CRYHO	Q5CGM9 cryptospori
748	16	64.0	363	2	Q8H1T5_ABUTH	Q8H1T5 abutilon th	821	16	64.0	428	2	Q6OXZ1_CABBR	Q6OXZ1 caenorhabdi
749	16	64.0	364	2	Q57WZ6_9TRYP	Q57WZ6 trypanosoma	822	16	64.0	429	2	Q4SUX3_TERNG	Q4SUX3 tetradon n
750	16	64.0	364	2	Q84XA2_MALXI	Q84XA2 malus xiao	823	16	64.0	432	2	Q8WZS7_NEUCR	Q8WZS7 neurospora
751	16	64.0	366	2	Q5NAP1_ORYSA	Q5NAP1 oryza sativ	824	16	64.0	433	2	Q7XLI3_ORYSA	Q7XLI3 oryza sativ
752	16	64.0	366	2	Q4TW98_CORIK	Q4TW98 corynebacte	825	16	64.0	433	2	Q8LFF0_ARATH	Q8LFF0 arabidopsis
753	16	64.0	366	2	Q4RJ83_TERNG	Q4RJ83 tetradon n	826	16	64.0	433	2	Q9ZYU4_ARATH	Q9ZYU4 arabidopsis
754	16	64.0	369	2	Q6G4L0_BARRB	Q6G4L0 bartonella	827	16	64.0	433	2	Q6BPN1_DEBHA	Q6BPN1 debaryomyce
755	16	64.0	370	2	Q7QRT3_GIALA	Q7QRT3 giardia lam	828	16	64.0	434	1	FXL20_HUMAN	FXL20 homo sapien
756	16	64.0	371	2	Q4LJTB_9BURK	Q4LJTB burkholderi	829	16	64.0	436	1	FXL20_MOUSE	FXL20 mus musculu
757	16	64.0	372	1	Q9STR6_DROME	Q9STR6 drosophila	830	16	64.0	436	2	Q8LIUT8_HUMAN	Q8LIUT8 homo sapien
758	16	64.0	377	1	BSC12_RAT	Q5FVJ6 rattus norv	831	16	64.0	437	2	Q5GRT08_BRARE	Q5GRT08 brachydanio
759	16	64.0	377	2	Q6ZRM1_HUMAN	Q6ZRM1 homo sapien	832	16	64.0	437	2	Q4SIUT7_TERNG	Q4SIUT7 tetradon n
760	16	64.0	377	2	Q9SUD8_ARATH	Q9SUD8 arabidopsis	833	16	64.0	437	2	Q9VSA4_DROME	Q9VSA4 drosophila
761	16	64.0	378	2	Q8OUV8_GVIRU	Q8OUV8 infectious	834	16	64.0	438	2		



835	16	64.0	438	2	Q5BDG6_BOVIN	Q58d96_bos_taurus	908	16	64.0	497	2	Q66B03_BABBO	Q68f03_babesia_bov
836	16	64.0	438	2	Q9DEV0_BRAAE	Q9dfv0_brachyadanto	909	16	64.0	499	2	Q4FKU6_9TRYP	Q4fkf6_rypanosoma
837	16	64.0	439	2	Q97DH7_CLOAB	Q97dh7_clostridium	910	16	64.0	500	2	Q20946_CAEEL	Q20946_caenornhabdi
838	16	64.0	441	2	Q4O7H1_LEIMA	Q4o7h1_leishmania	911	16	64.0	500	2	Q4Z0R8_PLABE	Q4z0r8_plasmodium
839	16	64.0	443	2	Q6EN36_ORYSA	Q6en36_oryza_sativ	912	16	64.0	502	2	Q8CHY5_MOUSE	Q8chy5_mus_musculu
840	16	64.0	445	2	Q818W6_GIALA	Q818w6_giardia_lam	913	16	64.0	504	1	BIT_F0GRU	Q8aby8_fuga_rubrip
841	16	64.0	446	2	Q6FM66_CANGA	Q6fm66_candida_gla	914	16	64.0	504	2	Q7SECI_NEUCR	Q7seci_neutrospora
842	16	64.0	446	2	Q84S01_ORYSA	Q84s01_oryza_sativ	915	16	64.0	504	2	Q6U118_9CAUD	Q6j118_bacterioph
843	16	64.0	447	2	Q4ODU2_LEIMA	Q4odu2_leishmania	916	16	64.0	505	2	Q7RR38_PLAYO	Q7rr38_plasmodium
844	16	64.0	449	2	Q8CFV1_XENIA	Q8cfv1_xenopus_lae	917	16	64.0	505	2	Q81756_9ATIC	Q81756_plasmodium
845	16	64.0	451	2	Q4PBR2_USTMA	Q4pbr2_ustlago_ma	918	16	64.0	506	2	Q571P7_MOUSE	Q571f7_mus_musculu
846	16	64.0	451	2	Q5TWL2_ANOCA	Q5twl2_anopheles_g	919	16	64.0	510	2	Q968M3_9TRYP	Q968m3_rypanosoma
847	16	64.0	451	2	Q4RML6_TETNG	Q4rml6_tetradon_n	920	16	64.0	510	2	Q583L3_9TRYP	Q583l3_rypanosoma
848	16	64.0	453	2	Q4UOZ6_9BUXK	Q4uoz6_burholderi	921	16	64.0	512	1	Q27072_TAEBO	Q27072_taeenia_soi
849	16	64.0	453	2	Q6A0E7_MOUSE	Q6a0e7_mus_musculu	922	16	64.0	512	1	VNN1_MOUSE	Q920k8_mus_musculu
850	16	64.0	454	2	Q81UW6_HUMAN	Q81uw6_homo_sapien	923	16	64.0	513	1	INAR2_MOUSE	Q571f7_mus_musculu
851	16	64.0	454	2	Q7R3V9_GIALA	Q7r3v9_giardia_lam	924	16	64.0	513	2	Q4FY6E_LEIMA	Q4fy6e_leishmania
852	16	64.0	456	2	Q25242_LUCCU	Q25242_lucilia_cup	925	16	64.0	513	2	Q923Z5_MOUSE	Q923z5_mus_musculu
853	16	64.0	457	2	Q7XPR7_ORYSA	Q7xpr7_oryza_sativ	926	16	64.0	515	2	Q4PD21_USTMA	Q4pd21_ustlago_ma
854	16	64.0	461	2	Q8A2G8_BACTN	Q8a2g8_bacteroides	927	16	64.0	515	2	Q708Y5_DROVI	Q708y5_drosophila
855	16	64.0	461	2	Q5BGE1_BRAAE	Q5bge1_brachyadanto	928	16	64.0	516	2	Q57ZFA_9TRYP	Q57zfa_rypanosoma
856	16	64.0	462	2	Q7Z8U4_LASPOR	Q7z8u4_aspergillus	929	16	64.0	517	2	Q7YPA2_CHLPH	Q7yep2_chlamydia_p
857	16	64.0	462	2	Q9C236_NEUCR	Q9c236_neutrospora	930	16	64.0	518	2	Q8N501_HUMAN	Q8n501_homo_sapien
858	16	64.0	463	2	Q5LIU4_BACFN	Q5liu4_bacteroides	931	16	64.0	518	2	Q4OUB7_THEAN	Q4oub7_chellieria_a
859	16	64.0	463	2	Q64ZY8_BACFR	Q64zy8_bacteroides	932	16	64.0	518	2	Q7D6T0_MYCTU	Q7d6t0_mycobacteri
860	16	64.0	465	2	Q4RM22_TETNG	Q4rm22_tetradon_n	933	16	64.0	520	2	Q4FKK1_9TRYP	Q4fkki_rypanosoma
861	16	64.0	466	2	Q757P9_ASHGO	Q757p9_ashya_goss	934	16	64.0	521	2	Q24867_HELRY	Q24867_helicobacte
862	16	64.0	466	2	Q7R7A2_PLAYO	Q7r7a2_plasmodium	935	16	64.0	523	2	Q7SYB1_BRAAE	Q7syb1_brachyadanto
863	16	64.0	467	2	Q5R9K3_FONPY	Q5r9k3_pongo_pygma	936	16	64.0	523	1	Y487_TREPA	Q83300_treponema_p
864	16	64.0	467	2	Q82L86_STRAW	Q82l86_streptomyce	937	16	64.0	525	1	Q7G9J6_ANODA	Q7g9j6_anopheles_g
865	16	64.0	468	1	Q53Y72_HUMAN	Q53y72_homo_sapien	938	16	64.0	526	2	Q6AOL6_DESFP	Q6aol6_desulfotale
866	16	64.0	468	2	Q9BER2_MACPA	Q9ber2_macca_fasc	939	16	64.0	528	2	P71969_MYCTU	Q7v552_mycobacteri
867	16	64.0	468	2	Q4RSX7_MACPA	Q4rsx7_macca_fasc	940	16	64.0	528	2	Q7TVS2_MYCBO	Q7v552_mycobacteri
868	16	64.0	468	2	Q6GV12_9INPA	Q6gv12_influenza_a	941	16	64.0	530	2	Q4IHU1_GIBB8	Q4ihu1_gibberella
869	16	64.0	469	2	Q6XUD6_IANNA	Q6xud6_influenza_a	942	16	64.0	530	2	Q9VWZ8_DROME	Q9vwz8_drosophila
870	16	64.0	469	2	Q6XV53_9INPA	Q6xv53_influenza_a	943	16	64.0	531	2	Q8PB46_XANCP	Q8pb46_xanthomonas
871	16	64.0	469	2	Q6XV61_9INPA	Q6xv61_influenza_a	944	16	64.0	532	2	Q94G13_PHYFA	Q94g13_pseudocitre
872	16	64.0	471	2	Q9VMG7_DROME	Q9vmg7_drosophila	945	16	64.0	533	2	Q7ZAP2_DESHP	Q7zap2_desulfovibr
873	16	64.0	472	2	Q83MW3_TROMT	Q83mw3_tropheryma	946	16	64.0	535	2	Q5CYB9_CRYPV	Q5cyb9_cryptospori
874	16	64.0	472	2	Q83NLO_TROM8	Q83nlo_tropheryma	947	16	64.0	535	2	Q4KKD6_PSEFP	Q4kkd6_pseudomonas
875	16	64.0	476	1	VTDB_RABIT	P53789_oryctolagus	948	16	64.0	536	2	Q27005_TOXDO	Q27005_coxoplasma
876	16	64.0	476	1	Q7Q250_GIALA	Q7q250_giardia_lam	949	16	64.0	537	2	Q5L6L4_CHLAP	Q5l6l4_chlamydochl
877	16	64.0	478	2	Q7OPAS_9DELT	Q7opas_melittangin	950	16	64.0	538	2	Q4R3F5_MACFA	Q4r3f5_macca_fasc
878	16	64.0	478	2	Q61W44_CAEER	Q61w44_caenornhabdi	951	16	64.0	539	2	Q4S0Z5_TETNG	Q4s0z5_tetradon_n
879	16	64.0	479	2	Q61W44_CAEER	Q61w44_caenornhabdi	952	16	64.0	540	1	ATG322_CANGA	Q6fy92_candida_gla
880	16	64.0	479	2	Q821U7_STRAE	Q821u7_streptomyce	953	16	64.0	541	1	Q5SNV1_CRYNE	Q5snv1_cryptococcu
881	16	64.0	480	2	Q60YB4_CAEER	Q60yb4_caenornhabdi	954	16	64.0	542	2	Q5KELO_CRYNE	Q5kelo_cryptococcu
882	16	64.0	480	2	Q25241_LUCCU	Q25241_lucilia_cup	955	16	64.0	542	2	Q8N7C2_HUMAN	Q8n7c2_homo_sapien
883	16	64.0	480	2	Q9X7G8_CAEEL	Q9x7g8_caenornhabdi	956	16	64.0	542	2	Q96192_HUMAN	Q96192_homo_sapien
884	16	64.0	481	2	Q4RKF0_TETNG	Q4rkf0_tetradon_n	957	16	64.0	542	2	Q92864_CHLPH	Q92864_chlamydia_p
885	16	64.0	482	2	Q5FM02_XENTR	Q5fm02_xenopus_tro	958	16	64.0	544	2	Q5B195_DROME	Q5b195_drosophila
886	16	64.0	482	2	Q800K7_PAROL	Q800k7_parallelity	959	16	64.0	545	1	TREA_PSEAE	Q91165_pseudomonas
887	16	64.0	483	2	Q4Z5Q4_PLABE	Q4z5q4_plasmodium	960	16	64.0	545	1	Q823Z7_CHLGV	Q823z7_chlamydochl
888	16	64.0	484	2	Q9M6F5_CHICK	Q9m6f5_gallus_gall	961	16	64.0	548	2	Q4FKD2_9TRYP	Q4fkfd2_rypanosoma
889	16	64.0	486	2	Q64AN8_9ARCH	Q64an8_uncultured	962	16	64.0	550	2	Q8RLB6_BRAJA	Q8rlb6_birdyfinzib
890	16	64.0	486	2	Q64AN8_9ARCH	Q64an8_uncultured	963	16	64.0	550	2	Q61G03_CAEER	Q61g03_caenornhabdi
891	16	64.0	486	2	Q4RU74_TETNG	Q4ru74_tetradon_n	964	16	64.0	551	2	Q7R3Y0_GIALA	Q7r3y0_giardia_lam
892	16	64.0	487	2	Q4Q9C4_LEIMA	Q4q9c4_leishmania	965	16	64.0	551	2	Q606V8_METCA	Q606v8_methylococc
893	16	64.0	487	2	Q8J1U8_XENIA	Q8j1u8_xenopus_lae	966	16	64.0	555	2	Q88CU1_PSEPK	Q88cu1_pseudomonas
894	16	64.0	488	2	Q64A02_CANGA	Q64a02_uncultured	967	16	64.0	557	1	GPEC1_HUMAN	Q35053_homo_sapien
895	16	64.0	489	2	Q6FWD4_CANGA	Q6fwd4_candida_gla	968	16	64.0	558	1	GPEC1_HUMAN	Q35053_homo_sapien
896	16	64.0	490	2	Q6IDV7_DROPS	Q6idv7_drosophila	969	16	64.0	558	1	Q53OM4_HUMAN	Q53om4_homo_sapien
897	16	64.0	492	1	Q4SSM9_TETNG	Q4ssm9_tetradon_n	970	16	64.0	558	2	Q6P7Q2_RAT	Q6p7q2_rattus_novr
898	16	64.0	492	1	TISD_HUMAN	P47974_homo_sapien	971	16	64.0	558	2	Q8HZ16_MACFA	Q8hz16_macca_fasc
899	16	64.0	493	1	Q50NG3_ENTHI	Q50ng3_entamoeba_h	972	16	64.0	559	2	Q5KEL1_CRYNE	Q5kel1_cryptococcu
900	16	64.0	493	1	LRC14_HUMAN	Q15048_homo_sapien	973	16	64.0	561	2	Q27007_TOXGO	Q27007_coxoplasma
901	16	64.0	493	2	Q57X17_9TRYP	Q57x17_rypanosoma	974	16	64.0	561	2	Q580T6_9TRYP	Q580t6_rypanosoma
902	16	64.0	493	2	Q569B5_RAT	Q569b5_rattus_novr	975	16	64.0	566	2	Q74BC4_GROSL	Q74bc4_greobacter_s
903	16	64.0	493	2	Q8VC16_MOUSE	Q8vc16_m_leucine_r	976	16	64.0	566	2	GAGI_DROME	P21330_drosophila
904	16	64.0	494	2	Q53TB4_HUMAN	Q53tb4_homo_sapien	977	16	64.0	568	1	Q68CN3_HUMAN	Q68cn3_homo_sapien
905	16	64.0	494	2	Q80WK3_MOUSE	Q80wk3_mus_musculu	978	16	64.0	571	1	Q6H9L7_HUMAN	Q6h9l7_homo_sapien
906	16	64.0	495	2	Q6MZD6_9ARCH	Q6mzd6_uncultured	979	16	64.0	571	2	Q8NBLO_HUMAN	Q8nblo_homo_sapien
907	16	64.0	496	2	Q66671_9GAMA	Q66671_equid_herpe	980	16	64.0	571	2		

```
981 16 64.0 571 2 Q6QH6_MOUSE Q6QH6_MOUSE
982 16 64.0 571 2 Q5EG07_ICPU Q5EG07_ICPU
983 16 64.0 572 2 Q5J8B6_9NCTU Q5J8B6_spodoptera
984 16 64.0 572 2 Q828T3_STRAW Q828T3_streptomyce
985 16 64.0 575 2 Q7QNV8_GITALA Q7QNV8_giardia lam
986 16 64.0 575 2 Q7XNV7_ORISA Q7XNV7_oryza sativ
987 16 64.0 577 2 Q7UX93_RHOBA Q7UX93_rhodospirill
988 16 64.0 578 2 Q7SG68_NEUCR Q7SG68_neutrospora
989 16 64.0 581 2 Q96MM9_BOTCI Q96MM9_botrytis ci
990 16 64.0 582 2 Q5J329_ORISA Q5J329_oryza sativ
991 16 64.0 587 1 SEL10_CABEL Q93794_caenorhabdi
992 16 64.0 588 2 Q6GUE9_TOBAC Q6GUE9_nicotiana t
993 16 64.0 588 2 Q5LXK6_BACPN Q5LXK6_bacteroides
994 16 64.0 588 2 Q64XS9_BACPN Q64XS9_bacteroides
995 16 64.0 592 2 Q60018_PICAN Q60018_pichia anqu
996 16 64.0 592 2 Q98H21_RHITO Q98H21_rhizobium l
997 16 64.0 595 2 Q9C2K3_NEUCR Q9C2K3_neutrospora
998 16 64.0 595 2 Q5VXL7_HUMAN Q5VXL7_homo sapien
999 16 64.0 596 1 HMDH1_SOLITU P48020_solanum tub
1000 16 64.0 596 2 Q4K8Z7_PSEPS Q4K8Z7_pseudomonas
```

## ALIGNMENTS

RESULT 1  
Q6IKH6\_DROME PRELIMINARY; PRT; 144 AA.

```
AC Q6IKH6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE HDCl2427.
GN ORFNames=HDC12427;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;
RA Hild M., Beckmann B., Haas S.A., Koch B., Solovyev V., Busold C.,
RA Fellenberg K., Boutros M., Vangron M., Sauer F., Hobeisel J.D.,
RA Paro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome."
RC Genome Biol. 5:RESEARCH0003.1-RESEARCH0003.17(2003).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK002390; DAA03896.1; -; Genomic DNA.
SQ SEQUENCE 144 AA; 15277 MW; 1E4588B8C183D5F2 CRC64;
```

Query Match 72.0%; Score 18; DB 2; Length 144;  
Best Local Similarity 22.2%; Pred. No. 27;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
DB 37 CATTASASC 45

RESULT 2  
P91214\_CABEL PRELIMINARY; PRT; 155 AA.

```
AC P91214;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE C-type lectin protein 45.
GN Name=clec-45; ORFNames=F07C4.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
```

OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxId=6239;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=Bristol N2; PubMed=9851916;  
RX MEDLINE=99069613; PubMed=9851916;  
RG The C. elegans sequencing consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology.";  
RT Science 282:2012-2018 (1998).  
RL EMBL; U80023; AAC24036.1; -; Genomic DNA.

DR PIR; T28945; T28945.  
DR Ensembl; F07C4.2; Caenorhabditis elegans.  
DR Wormbase; WBGene0017199; F07C4.2.  
DR WormPep; F07C4.2; CB09201.  
DR GO; GO:0005529; F-sugar binding; IEA.  
DR InterPro; IPR01304; Lectin\_C.  
DR SMART; SM00034; CLECT; 1.  
DR PROSITE; PS00615; C TYPE LECTIN 1; UNKNOWN\_1.  
DR PROSITE; PS50041; C TYPE LECTIN\_2; 1.  
KW Complete proteome; Lectin.  
SQ SEQUENCE 155 AA; 16285 MW; 956F0F4FDCB0F93 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 155;  
Best Local Similarity 22.2%; Pred. No. 28;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
DB 92 CTTASSSSC 100

## RESULT 3

```
Q7PRQ7_ANOGA PRELIMINARY; PRT; 269 AA.
AC Q7PRQ7;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000001657 (Fragment).
GN ORFNames=ENSANGG0000001387;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxId=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008847; BAA06779.3; -; Genomic DNA.
FT NON_TER 1
FT TER 269
SQ SEQUENCE 269 AA; 24228 MW; 25B8F71FD71F1F2 CRC64;
```

Query Match 72.0%; Score 18; DB 2; Length 269;  
Best Local Similarity 22.2%; Pred. No. 34;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9  
DB 228 C8SSTSSC 236

## RESULT 4

O6GZ05\_MAIZE PRELIMINARY; PRT; 281 AA.

AC O6GZ05;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE DREB-like protein.

GN Name=drebl;

OS Zea mays (Maize);

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC PACCAD clade; Panicoideae; Andropogoneae; Zea.

NCBI\_TaxID=4577;

RN NUCLEOTIDE SEQUENCE.

RA Qiu F., Li J., Zhao J., Chen S.-Y., Liu O.;

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF48789; AAN7673.1; -; mRNA.

DR HSSP: O80337; 2GCC.

DR GO: GO:0005634; C:nucleus; IEA.

DR GO: GO:0003700; F:transcription factor activity; IEA.

DR InterPro: IPR001471; TF\_ERF.

DR Pfam: PF00847; AP2; 1.

DR PRINTS: PR00367; ETHRSPLEMT.

DR PRODOM: PD001423; TF\_ERF; 1.

DR SMART: SM00380; AP2; 1.

SQ SEQUENCE 281 AA; 29346 MW; A0902A7FE0AB978 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 281;  
Best Local Similarity 22.2%; Pred. No. 35;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 CXXXXXXC 9  
DB 11 CSATTSSC 19

RESULT 5  
O6LKK0\_MAIZE PRELIMINARY; PRT; 281 AA.

AC O6LKK0;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE DRE binding factor 2.

GN Name=dre2;

OS Zea mays (Maize);

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC PACCAD clade; Panicoideae; Andropogoneae; Zea.

NCBI\_TaxID=4577;

RN NUCLEOTIDE SEQUENCE.

RA MEDLINE=22058745; PubMed=12061899;

RA Kizis D., Pages M.;

RT "Maize DRE-binding proteins DBF1 and DBF2 are involved in rab17

RT regulation through the drought-responsive element in an ABA-dependent

RT pathway.";

RL Plant J. 30:679-689 (2002).

DR EMBL: AF493799; AAM80485.1; -; mRNA.

DR HSSP: O80337; 2GCC.

DR GO: GO:0005634; C:nucleus; IEA.

DR GO: GO:0003700; F:transcription factor activity; IEA.

DR InterPro: IPR001471; TF\_ERF.

DR Pfam: PF00847; AP2; 1.

DR PRINTS: PR00367; ETHRSPLEMT.

DR PRODOM: PD001423; TF\_ERF; 1.

DR SMART: SM00380; AP2; 1.

SQ SEQUENCE 281 AA; 29332 MW; CBE02A7FE0AB978 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 281;  
Best Local Similarity 22.2%; Pred. No. 35;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 CXXXXXXC 9  
DB 11 CSATTSSC 19

RESULT 6  
O80S50\_9BETA PRELIMINARY; PRT; 347 AA.

AC O80S50;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE UL38.

OS Pongine herpesvirus 4 (Chimpanzee cytomegalovirus).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Cytomegalovirus.

NCBI\_TaxID=188763;

RN NUCLEOTIDE SEQUENCE.

RA MEDLINE=22421467; PubMed=12533697; DOI=10.1099/vir.0.18606-0;

RA Davidson A.J., Dolan A., Akter P., Addison C., Dargan D.J.,

RA Alcindor D.J., McGeoch D.J., Hayward G.S.;

RT "The human cytomegalovirus genome revisited: comparison with the

RT chimpanzee cytomegalovirus genome.";

RT J. Gen. Virol. 84:17-28 (2003).

DR EMBL: AF480884; AAM00688.1; -; Genomic DNA.

SQ SEQUENCE 347 AA; 39141 MW; 34B18107D3501982 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 347;  
Best Local Similarity 22.2%; Pred. No. 38;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 CXXXXXXC 9  
DB 312 CSSTSSSTC 320

RESULT 7  
O9ST87\_ORYSA PRELIMINARY; PRT; 416 AA.

AC O9ST87;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE CAA303710.1 protein.

GN Name=q3037.10;

OS Oryza sativa (rice);

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaeae; Oryza.

NCBI\_TaxID=4530;

RN NUCLEOTIDE SEQUENCE.

RA STRAIN=DNA;

RA Hong G., Chen Z.;

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ245900; CAB53483.1; -; Genomic DNA.

DR Gramene: O9ST87; -;

DR GO: GO:0004629; F:phospholipase C activity; IEA.

DR GO: GO:0007242; P:intracellular signaling cascade; IEA.

DR InterPro: IPR000909; PL\_PLC\_X.

DR PROSITE: PS50007; PIP2C\_X DOMAIN; 1.

SQ SEQUENCE 416 AA; 45736 MW; 332A261949FBE6D CRC64;

Query Match 72.0%; Score 18; DB 2; Length 416;  
Best Local Similarity 22.2%; Pred. No. 41;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 DB 34 CTASASASC 42

RESULT 8  
 QXST8 ORYSA PRELIMINARY; PRT; 468 AA.  
 AC QXST8;  
 DT 01-OCT-2003 (TReMBLrel. 25, Created)  
 DT 01-MAR-2004 (TReMBLrel. 26, last sequence update)  
 DE OSJBA0039K24.18 protein.  
 GN Name=OSJBA0039K24.18;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OC NCBI\_TaxID=39947;  
 RN [1]

RESULT 9  
 Q9LFB4 ARATH PRELIMINARY; PRT; 470 AA.  
 AC Q9LFB4;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)  
 DE Hypothetical protein F7J8\_200.  
 GN Name=F7J8\_200;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]

RESULT 10  
 Q9XZX8 LEIMA PRELIMINARY; PRT; 504 AA.  
 AC Q9XZX8;  
 DT 01-NOV-1999 (TReMBLrel. 12, Created)  
 DT 01-NOV-1999 (TReMBLrel. 12, last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, last annotation update)  
 DE Possible surface antigen.  
 GN Name=L302.01;  
 OS Leishmania major.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OC NCBI\_TaxID=5664;  
 RN [1]

RESULT 11  
 Q8S4F6 ARATH PRELIMINARY; PRT; 510 AA.  
 AC Q8S4F6;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, last sequence update)  
 DT 01-FEB-2005 (TReMBLrel. 29, last annotation update)

RP NUCLEOTIDE SEQUENCE.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL131189; CAB69850.1; -; Genomic\_DNA.  
 DR PIR; T45962; T45962.  
 DR GO; GO:0009058; P:biogenesis; IEA.  
 DR InterPro; IPR001296; Glyco trans 1.  
 DR Pfam; PF00534; Glyco\_transf\_1; I.  
 KW Hypothetical protein.  
 SQ SEQUENCE 470 AA; 5197 MW; A7CBBBDABEF5BC6 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 470;  
 Best Local Similarity 22.2%; Pred. No. 43;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 DB 161 CASTASTSC 169

RESULT 12  
 Q8S4F6 ARATH PRELIMINARY; PRT; 510 AA.  
 AC Q8S4F6;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, last sequence update)  
 DT 01-FEB-2005 (TReMBLrel. 29, last annotation update)

Query Match 72.0%; Score 18; DB 2; Length 504;  
 Best Local Similarity 22.2%; Pred. No. 44;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 DB 34 CTASASASC 42

RESULT 13  
 Q9LFB4 ARATH PRELIMINARY; PRT; 470 AA.  
 AC Q9LFB4;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)  
 DE Hypothetical protein F7J8\_200.  
 GN Name=F7J8\_200;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]

RESULT 14  
 Q9XZX8 LEIMA PRELIMINARY; PRT; 504 AA.  
 AC Q9XZX8;  
 DT 01-NOV-1999 (TReMBLrel. 12, Created)  
 DT 01-NOV-1999 (TReMBLrel. 12, last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, last annotation update)  
 DE Possible surface antigen.  
 GN Name=L302.01;  
 OS Leishmania major.  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OC NCBI\_TaxID=5664;  
 RN [1]

RESULT 15  
 Q8S4F6 ARATH PRELIMINARY; PRT; 510 AA.  
 AC Q8S4F6;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, last sequence update)  
 DT 01-FEB-2005 (TReMBLrel. 29, last annotation update)

DE Sulfolipid synthase [Hypothetical protein At5g01220].  
GN Name=SQD2; Synonyms=AC5901220;  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=11960029; DOI=10.1073/pnas.082696499;  
RT Yu B., Xu C., Benning C.,  
RT "Arabidopsis disrupted in SQD2 encoding sulfolipid synthase is  
RT impaired in phosphate-limited growth."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:5732-5737(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,  
RA Onodera C.S., Quach H.L., Tang C.C., Toriumi M., Wong C., Wu H.C.,  
RA Yu G., Yuan S., Chen H., Cheuk R., Jones T., Kim C.J., Nguyen M.,  
RA Palm C.J., Shim P., Southwick A., Tripp M.G., Wu T., Davis R.W.,  
RA Becker J.R., Theologis A.;  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF454354; AAM18913.1; -; mRNA.  
DR EMBL; BT005796; AAO64198.1; -; mRNA.  
DR GO; GO:0009058; P:bioynthesis; IEA.  
DR InterPro; IPR001296; Glyco\_transf\_1.  
DR Pfam; PF00534; Glycos\_transf\_1; 1.  
KW Hypothetical protein  
SQ SEQUENCE 510 AA; 56630 MW; 20E57D318DBED68 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 510;  
Best Local Similarity 22.2%; Pred. No. 44;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 23 CXXXXXXC 31

RESULT 12  
ID 07X21\_XENIA PRELIMINARY; PRT; 513 AA.  
AC 07X21;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE MG35520 protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
OC Xenopodinae; Xenopus; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenn C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative."  
RL Dev. Dyn. 225:384-391(2002).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Embryo;  
RA Klein S., Strausberg R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC046254; AAM46254.1; -; mRNA.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR003323; OTU.  
DR Pfam; PF02338; OTU; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS00802; OTU; 1.  
SQ SEQUENCE 513 AA; 55868 MW; A8734AD686FEEDA CRC64;

Query Match 72.0%; Score 18; DB 2; Length 513;  
Best Local Similarity 22.2%; Pred. No. 44;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 376 CXXXXXXC 384

RESULT 13  
ID 0640H3\_XENIA PRELIMINARY; PRT; 518 AA.  
AC 0640H3;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE LOC494714 protein.  
GN Name=LOC494714;  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
OC Xenopodinae; Xenopus; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Kidney;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative."  
RL Dev. Dyn. 225:384-391(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Kidney;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenn C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.M., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smilun D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RN NUCLEOTIDE SEQUENCE.  
 RP TISSUE=Kidney;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC082654; AAH82654.1; -; mRNA.  
 DR InterPro: IPR003323; OTU.  
 DR Pfam: PF02338; OTU; 1.  
 DR PROSITE: PS50802; OTU; 1.  
 SQ SEQUENCE 518 AA; 56209 MW; 2099AB07C6C93EC1 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 518;  
 Best Local Similarity 22.2%; Pred. No. 45;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 9  
 Db 381 CSSTAAAC 389

RESULT 14  
 Q70XT3 GIALA PRELIMINARY; PRT; 569 AA.  
 AC Q70XT3;  
 DT 01-MAR-2004 (TREMBLrel. 26, Created)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)  
 DE GIP 77 40692 38983.  
 OS Giardia lamblia ATCC 50803.  
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
 NCBI\_TaxID=184922;  
 RN [1]  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=WB C6;  
 RA MEDLINE=21167845; PubMed=11104758; DOI=10.1074/jbc.M006589200;  
 RA Abel E.S., Davies B.J., Robles L.D., Loflin C.E., Gillin F.D.,  
 RA Chakrabarti R.;  
 RT "Possible roles of protein kinase A in cell motility and excystation  
 RT of the early diverging eukaryote Giardia lamblia.";  
 RL J. Biol. Chem. 276:10320-10329(2001).  
 RN [2]  
 RN NUCLEOTIDE SEQUENCE.  
 RP STRAIN=WB C6;  
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,  
 RA Olsen G.J., Sogin M.L.;  
 RT "Draft sequence of the Giardia lamblia genome.";  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC Preliminary data.  
 DR EMBL: AACB01000056; EAA39878.1; -; Genomic DNA.  
 DR GO: GO:0005489; F:electron transporter activity; IEA.  
 DR GO: GO:0005506; F:iron ion binding; IEA.  
 DR GO: GO:0006118; P:electron transport; IEA.  
 DR InterPro: IPR001450; Pfam: PF03302; VSP; 1.  
 DR InterPro: IPR005127; Giardia\_VSP.  
 DR Pfam: PF03302; VSP; 1.  
 DR PRINTS: PR00353; 4FE4SFRDOXIN.  
 KM Electron transport; Transport.  
 SQ SEQUENCE 569 AA; 59086 MW; EB0DF4E2B8CC3F20 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 569;  
 Best Local Similarity 22.2%; Pred. No. 46;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 9  
 Db 464 CSSTAAAC 472

RESULT 15  
 Q7XV79 ORYSA PRELIMINARY; PRT; 684 AA.  
 AC Q7XV79;  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)  
 DE OSJNB0076A22.6 protein.  
 GN Name=OSJNB0076A22.6;  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 NCBI\_TaxID=4530;  
 RN [1]  
 RN NUCLEOTIDE SEQUENCE.  
 RP MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;  
 RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,  
 RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,  
 RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,  
 RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Gan J.,  
 RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yan H.,  
 RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,  
 RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,  
 RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,  
 RA Lan L., Lai Y., Cheng Z., Gu W., Jiang D., Li J., Hong G., Xue Y.,  
 RT "Sequence and analysis of rice chromosome 4.";  
 RL Nature 420:316-320(2002).  
 DR EMBL: AL663016; CAD40795.3; -; Genomic DNA.  
 DR Gramene: Q7XV79;  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; Pkinase; 1.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_SF; UNKNOWN 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_SF; UNKNOWN 1.  
 SQ SEQUENCE 684 AA; 75877 MW; 8CB1910F10409B14 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 684;  
 Best Local Similarity 22.2%; Pred. No. 50;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 9  
 Db 20 CSAAATATC 28

RESULT 16  
 DLA\_BRAE STANDARD; PRT; 772 AA.  
 ID DLA\_BRAE  
 AC Q6D1A8; O57462;  
 DT 13-SEP-2005 (Rel. 48, Created)  
 DT 13-SEP-2005 (Rel. 48, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Delta-like protein A precursor (DeltaA protein).  
 GN Name=dlp;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 CC Cyprinidae; Danio.  
 CC NCBI\_TaxID=7955;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, AND DEVELOPMENTAL  
 STAGE.  
 RX MEDLINE=98165392; PubMed=9425133;  
 RA Appel B., Eiseen J.S.;  
 RT "Regulation of neuronal specification in the zebrafish spinal cord by  
 Delta function.";  
 RL Development 125:371-380(1998).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC TISSUE=Embryo;  
 RG NIH - Zebrafish Gene Collection (ZGC) project;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP FUNCTION.  
 RX PubMed=10572043;  
 RA Riley B.B., Chiang M.-Y., Farmer L., Heck R.;  
 RT "The delta gene of zebrafish mediates lateral inhibition of hair  
 cells in the inner ear and is regulated by pax2.1.";  
 RL Development 126:5669-5678(1999).  
 RN [4]  
 RP INTERACTION WITH MIB, AND UBIQUITINATION.  
 RX PubMed=15013799; DOI=10.1016/j.ydbio.2003.11.010;  
 RA Chen W., Corliss D.C.;  
 RT "Three modules of zebrafish Mind bomb work cooperatively to promote  
 Delta ubiquitination and endocytosis.";  
 RL Dev. Biol. 267:361-373(2004).  
 CC -1- FUNCTION: Acts as a ligand for Notch receptors and is involved in  
 primary neurogenesis. Can activate Notch receptors, thereby  
 playing a key role in lateral inhibition, a process that prevents  
 the immediate neighbors of each nascent neural cell from  
 simultaneously embarking on neural differentiation. In inner ear,  
 it prevents adjacent cells from adopting the same cell fate.  
 CC -1- SUBUNIT: Interacts with mib.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed in nervous system. In the developing  
 nervous system, it is expressed in overlapping regions with deltab  
 (dlt) and deltal (dlt), but differs the neural plate: it is  
 expressed in patches of contiguous cells with dlt, while dlt is  
 apparently confined to the scattered cells within those patches  
 that differentiate as neurons. Expressed in hair cells of inner  
 ear.  
 CC -1- DEVELOPMENTAL STAGE: Initiated in the neuroectoderm before that of  
 dlt. In the developing trunk neural plate and neural tube, it is  
 initiated in the epiblast prior to completion of gastrulation. At  
 the 2- to 3-somite stage (10.5 hours) low levels are distributed  
 throughout the trunk CNS, with cells expressing higher levels  
 found in the medial and lateral regions of the neural plate. These  
 regions correspond to the positions at which primary motoneurons  
 and Rohon Bead neurons (Rbns) originate. Cells expressing high  
 levels do not form contiguous domains. Rather, single cells or  
 small clusters of several cells showing high expression are  
 interspersed with cells having lower expression. Expressed is  
 specific to the developing nervous system, and continues to be  
 expressed broadly in the CNS throughout neurogenesis. Expressed in  
 cells specified for neuronal fates. At 24 hours, and throughout  
 later embryogenesis, it is broadly expressed in the spinal cord,  
 suggesting that it is expressed by many types of cells. Expressed  
 as neuronal specification occurs and is subsequently down-  
 regulated in cells that have acquired specific neuronal fates.  
 CC -1- PTM: Ubiquitinated by mib, leading to its endocytosis and  
 subsequent degradation.  
 CC -1- SIMILARITY: Contains 1 DSL domain.  
 CC -1- SIMILARITY: Contains 8 EGF-like domains.  
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a  
 frameshift in position 726.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL: AF030031; AAC41249.1; ALT SEQ; mRNA.  
 DR EMBL: BC075742; AAH75742.1; -, mRNA.  
 DR ZFIN: ZDB-GENE-980526-23; dlt.  
 DR InterPro: IPR000152; Asx\_hydroxyl\_S.  
 DR InterPro: IPR001774; DSL.  
 DR InterPro: IPR000742; EGF\_2.  
 DR InterPro: IPR001881; EGF\_Ca\_bd.  
 DR InterPro: IPR001438; EGF\_II.  
 DR InterPro: IPR006209; EGF-like.  
 DR InterPro: IPR006210; IEGF.  
 DR Pfam: PF01414; DSL\_1.  
 DR Pfam: PF00008; EGF\_6.  
 DR Pfam: PF07657; MNML\_1.  
 DR PRINTS: PR00010; EGFBLD.  
 DR SMART: SM00051; DSL\_1.  
 DR SMART: SM00181; EGF\_1.  
 DR SMART: SM00179; EGF\_CA; 6.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 2.  
 DR PROSITE: PS1051; DSL\_1.  
 DR PROSITE: PS00022; EGF\_1; 8.  
 DR PROSITE: PS01186; EGF\_2; 8.  
 DR PROSITE: PS50026; EGF\_3; 7.  
 DR PROSITE: PS01187; EGF\_CA; 1.  
 KW Calcium; Developmental protein; Differentiation; EGF-like domain;  
 KW Glycoprotein; Neurogenesis; Notch signaling pathway; Repeat; Signal;  
 KW Transmembrane; Ubl conjugation.  
 FT SIGNAL 1 20  
 FT CHAIN 1 772  
 FT TOPO\_DOM 21 536  
 FT TRANSMEM 537 557  
 FT TOPO\_DOM 558 772  
 FT DOMAIN 179 223  
 FT DOMAIN 225 257  
 FT DOMAIN 257 288  
 FT DOMAIN 290 330  
 FT DOMAIN 330 366  
 FT DOMAIN 366 405  
 FT DOMAIN 407 443  
 FT DOMAIN 445 481  
 FT DOMAIN 483 519  
 FT COMPIAS 697 746  
 FT CARBHYD 479 499  
 FT DISULFID 228 239  
 FT DISULFID 232 245  
 FT DISULFID 259 270  
 FT DISULFID 265 276  
 FT DISULFID 278 287  
 FT DISULFID 294 306  
 FT DISULFID 300 316  
 FT DISULFID 318 327  
 FT DISULFID 334 345  
 FT DISULFID 339 354  
 FT DISULFID 356 365  
 FT DISULFID 372 383  
 FT DISULFID 377 393  
 FT DISULFID 395 404  
 FT DISULFID 411 422  
 FT DISULFID 416 431  
 FT DISULFID 433 442  
 FT DISULFID 449 460  
 FT DISULFID 454 469  
 FT DISULFID 471 480  
 FT DISULFID 487 498  
 FT DISULFID 492 507  
 FT DISULFID 509 518  
 SO SEQUENCE 772 AA; 84969 MW; 716A014158938576 CRC64;  
 Query Match 72.0%; Score 18; DB 1; Length 772;  
 Best Local Similarity 22.2%; Pred. No. 52;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

DB 740 CASASTSAC 748

RESULT 17

Q4WY29 ASPFU PRELIMINARY; PRT; 828 AA.

AC Q4WY29;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Repressor protein.

GN ORFNames=AF3914830;

OS Aspergillus fumigatus Af293.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI\_TaxID=330879;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=AF293;

RA Nieman W., Pain A., Anderson M.J., Mortman J., Kim H., Stanley J.,

RA Arroya J., Berriman M., Abe K., Archer D.B., Bernedo C., Bennett J.,

RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,

RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,

RA Foster N., Fraser A., Garcia J.L., Garcia M., Godle A.,

RA Goldman G.H., Goni K., Griffith-Jones S., Gwilliam R., Haas B.,

RA Haas H., Harris D., Horluchi H., Huang J., Humphrey S., Jimenez J.,

RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,

RA Kumagai T., Latfon A., Latge J.-P., Li W., Lord A., Lu C.,

RA Majors W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,

RA Mouyia I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,

RA Penava M.A., Petrea M., Price C., Pritchard B.L., Quail M.A.,

RA Rabinowitsch E., Rawlins N., Rajandream M.-A., Reichard U.,

RA Renaldi H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,

RA Roming C.M., Ruter S., Salzberg S.L., Sanchez M.,

RA Sanchez-Ferrero J.C., Sanders D., Seeger K., Squares R., Squares S.,

RA Takeuchi M., Tekala F., Turner G., Vazquez de Alana C.R., Weidman J.,

RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,

RA Machida M., Hall N., Barrell B., Denning D.W.;

RT "Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus.";

RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

CC DR EMBL; AAHF0100002; EAL92104.1; -; Genomic DNA.

CC DR EMBL; AAHF0100002; 2B1B831C5B4F0C6 CRC64;

SQ SEQUENCE 828 AA; 92421 MW; 2B1B831C5B4F0C6 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 828;

Best Local Similarity 22.2%; Pred. No. 54;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

DB 57 CTTTTTTC 65

RESULT 18

Q4SR96 TETNG PRELIMINARY; PRT; 1013 AA.

AC Q4SR96;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Chromosome 11 SCAP14528, whole genome shotgun sequence.

DE (Fragment)

GN ORFNames=GSTENG00014015001;

OS Tetraodon nigroviridis (Green puffer).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;

OC Tetraodontidae; Tetraodontidae; Tetraodon.

OX NCBI\_TaxID=99883;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,

RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costat C., Bernot A.,

RA Nicard S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segreus B.,

RA Dasila C., Salanoubat M., Levy M., Boudet N., Castellano S.,

RA Anthouard V., Jabin C., Castell V., Katinka M., Vacherie B.,

RA Biemont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,

RA Craud C., Duprat S., Brothier P., Contancieu J.P., Guzy J.,

RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,

RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,

RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,

RA Laudet V., Schachter V., Quetier F., Sautin W., Scarpelli C.,

RA Wincker P., Lander E.S., Weissbach J., Roest Crollins H.;

RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals

RT the early vertebrate proto-karyotype.";

RL Nature 431:946-957(2004).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RG Genoscope, Whitehead Institute Centre for Genome Research;

RL Submitted (Feb-2004) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

CC DR EMBL; CAAB01014528; CAP96836.1; -; Genomic DNA.

DR Interpro; IPR000175; Na/nttran\_sympor.

DR Pfam; PF00209; SNE; 2.

DR PRINTS; PR00176; NANUSMPOR.

DR Prodom; PD000448; Na/nttran\_sympor; 2.

DR PROSITE; PS00610; NA\_NEUROTRAN\_SYMP\_1; 1.

DR PROSITE; PS00754; NA\_NEUROTRAN\_SYMP\_2; UNKNOWN\_1.

DR PROSITE; PS0267; NA\_NEUROTRAN\_SYMP\_3; 2.

KW Symport; Transmembrane; Transport.

FT NON TER 1

FT NON TER 1

SQ SEQUENCE 1013 AA; 110756 MW; 3D5SCFE7787652D CRC64;

Query Match 72.0%; Score 18; DB 2; Length 1013;

Best Local Similarity 22.2%; Pred. No. 58;

Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

DB 728 CSASTSAC 736

RESULT 19

Q4QFB5 LEIMA PRELIMINARY; PRT; 1031 AA.

AC Q4QFB5;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=LmjF15.0700;

OS Leishmania major.

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OC NCBI\_TaxID=5664;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Friddlin;

RA Peacock C.S., Murphy L., Ivans A.C., Berriman M., Blackwell J.,

RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,

RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;

RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.

DR EMBL; CTO05254; CAJ03294.1; -; Genomic DNA.

DE Hypothetical protein.

SQ SEQUENCE 1031 AA; 112673 MW; 65DB334CA5FCBA CRC64;

Query Match 72.0%; Score 18; DB 2; Length 1031;



Best Local Similarity 22.2%; Pred. No. 59;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 791 CXXXXSSSAC 799

RESULT 20  
O5CV70\_CRYPV PRELIMINARY; PRT; 1107 AA.

AC O5CV70;  
DT 10-MAY-2005 (TEMBUREL. 30, Created)  
DT 10-MAY-2005 (TEMBUREL. 30, Last sequence update)  
DE Hypothetical protein.  
GN ORFNames=cg48\_5210;  
OS Cryptosporidium parvum.  
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
OC Cryptosporidiidae; Cryptosporidium.  
OX NCBI\_TaxId=5807;

RP NUCLEOTIDE SEQUENCE.  
STRAIN=Iowa type II;  
RX PubMed=15044751; DOI=10.1126/science.1094786;  
RA Abrahamson M.S., Templeton T.J., Enomoto S., Abrahante J.B., Zhu G.,  
RA Lancio C.A., Deng M., Liu C., Widmer G., Tzipori S., Buck G.A., Xu P.,  
RA Bankier A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,  
RA Anantharaman V., Aravind L., Kapur V.;  
RT "Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*."  
RL Science 304:441-445 (2004).  
CC -!- CAUTION: The sequence shown here is derived from an  
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
preliminary data.

CC EMBL; AAE0100003; EAK89643.1; -; Genomic DNA.  
DR InterPro: IPR012283; 6PGD C core.  
DR InterPro: IPR011989; ARM-like.  
DR InterPro: IPR012292; Globin-related.  
DR InterPro: IPR001524; Glyco Hydro-6.  
DR InterPro: IPR012287; Homeodomain-rel.  
DR InterPro: IPR001093; IMPH/GMPase.  
DR InterPro: IPR008976; PLAT LH2.  
DR InterPro: IPR000783; RNAPol RPBS.  
DR InterPro: IPR009032; Vpu cyf.  
DR InterPro: IPR011981; Wing\_hlx\_DNA\_bd.  
DR Pfam: PF00478; IMPDH; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 1107 AA; 127343 MW; 4BAD9AE836ECB2AF CRC64;

Query Match 72.0%; Score 18; DB 2; Length 1107;  
Best Local Similarity 22.2%; Pred. No. 60;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9

Db 185 CXXXXSSSAC 193

RESULT 21  
E75BC DROME  
ID E75BC DROME STANDARD; PRT; 1199 AA.  
AC P167L; O810S1; O810S3;  
DT 01-ANG-1990 (Rel. 15, Created)  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Ecdysone-induced protein 75B isoforms C/D (E75-A).  
GN Name=E75B; Synonyms=NR1D3; ORFNames=CG8127;  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxId=7227;

RN [1]  
RC NUCLEOTIDE SEQUENCE [MRNA], AND ALTERNATIVE SPLICING.  
RP STRAIN=Cancon-S;  
RX MEDLINE=90249727; PubMed=2110921;  
RA Segaves W.A., Hognes D.S.;  
RT "The E75 ecdysone-inducible gene responsible for the 75B early puff in  
RT Drosophila encodes two new members of the steroid receptor  
RT superfamily."  
RL Genes Dev. 4:204-219 (1990).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Chao M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kelchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasako P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Pacleib J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Maestaran D.A., Weinstein D.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*."  
RL Science 287:2185-2195 (2000).  
RN [3]  
RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Betencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review."  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).  
RN [4]  
RP FUNCTION, DEVELOPMENTAL STAGE, AND INDUCTION.  
RX MEDLINE=94038699; PubMed=8223281;  
RA Huet F., Ruiz C., Richards G.;  
RT "Puffs and PCR: the in vivo dynamics of early gene expression during  
RT ecdysone responses in *Drosophila*."  
RL Development 118:613-627 (1993).  
CC -!- FUNCTION: Implicated in the regulation of ecdysone-triggered gene  
hierarchies. Probably plays a key role in mediating the regulation

```

CC of the larval molt by 20-OH-ecdysone.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=C; Synonyms=E75A;
CC IsoId=P1671-1; Sequence=Displayed;
CC Name=A; Synonyms=E75B;
CC IsoId=P1672-1; Sequence=External;
CC Name=B; Synonyms=E75C;
CC IsoId=P1055-1; Sequence=External;
CC Name=D;
CC IsoId=P1671-2; Sequence=VSP_014915, VSP_014916;
CC Note=No experimental confirmation available;
CC -1- DEVELOPMENTAL STAGE: In mid instar larvae salivary glands, low
CC basal levels are observed in puff stage 1. Levels increase in late
CC larvae from puff stages 3-10, then decrease abruptly at stage 11.
CC In prepupae, isoform C is the predominant form during the
CC transition between puff stages 18-19. At puff stage 1, expression
CC is also present in the gut. By stage 3 it is present in the wing
CC disks, Malpighian tubules and the fat body. At stage 11,
CC expression is only present in the gut and wing disks.
CC -1- INDUCTION: The expression of this protein is developmentally
CC regulated and is correlated with the 20-OH-ecdysone induced
CC activity of puff 75B.
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family. NR1
CC subfamily.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 1164.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL, X51548; CA35923.1; ALT FRAME; mRNA.
DR EMBL, AE003522; AAN11687.1; -; Genomic DNA.
DR EMBL, AE003522; AAN11689.1; -; Genomic DNA.
DR PIR, A34598; A34598.
DR HSSP, P20393; 1GAS.
DR SMR, P1671; 240-323.
DR TRANSFAC, T01367; -.
DR FLYBASE, FBgn0000568; E1p75B.
DR GO, GO:0004879; F.ligand-dependent nuclear receptor activity; NAS.
DR GO, GO:0018990; P.ecdysis (sennu insecta); IMP.
DR GO, GO:0035072; P.ecdysone-mediated induction of salivary gla. . .; NAS.
DR GO, GO:0007553; P.regulation of ecdysteroid metabolism; IMP.
DR InterPro, IPR001628; Hrmn_rcpt_DNA_bd.
DR InterPro, IPR000536; Hrmn_rcpt_lig_bd.
DR InterPro, IPR001723; Stdnrm_receptor.
DR InterPro, IPR000324; VltD_receptor.
DR Pfam, PF00104; Hormone_recep. 1.
DR Pfam, PF00105; zf-C4; 1.
DR PRINTS, PR00398; STRDHORMNER.
DR PRINTS, PR00447; STROIDFINGER.
DR PRINTS, PR00350; VITAMINDER.
DR PRODOM, PD0000035; znf_C4steroid; 1.
DR SMART, SM00430; HO1; 1.
DR SMART, SM00399; znf_C4; 1.
DR PROSITE, PS00031; NUCLEAR_REC_DBD_1; 1.
DR PROSITE, PS1030; NUCLEAR_REC_DBD_2; 1.
KW Alternative splicing; Developmental protein; DNA-binding;
KW Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
KW DNA_BIND 242 318 Nuclear receptor.
FT ZN_FING 245 265 NR C4-type.
FT ZN_FING 282 306 NR C4-type.
FT REGION 380 590 Ligand-binding (potential).
FT COMPIAS 139 195 Gln/His-rich.
FT COMPIAS 721 756 Gln-rich.
FT COMPIAS 900 1103 Ser-rich.
FT VARSPIC 1 293 Missing (in isoform D).

```

```

FT FT VARSPIC 294 315 /FTId=VSP_014915.
FT FT VARSPIC 294 315 NNRRCYCRILKICIVAGMSRDA -> MGEELPIILKILKGN
FT FT VARSPIC 294 315 VNYHNAP (in isoform D).
FT FT VARSPIC 294 315 /FTId=VSP_014916.
SQ SEQUENCE 1199 AA; 128534 MM; E29372F5E29F55 CRC64;
Query Match 72.0%; Score 18; DB 1; Length 1199;
Best local Similarity 22.2%; Pred. No. 62;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CXXXXXXX 9
Db 1070 CXXXXXXX 1078
RESULT 22
E75BA_DROME STANDARD; PRT; 1355 AA.
ID E75BA_DROME
AC P17672; Q810S2;
DT 01-AUG-1990 (Rel. 15, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Ecdysone-induced protein 75B isoform A (E75-B).
GN Name=E1p75B; Synonyms=NR1D3; ORFNames=CG8127;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND ALTERNATIVE SPLICING.
RC STRAIN=Canton-S;
RX MEDLINE=90249727; PubMed=2110921;
RT Segreaves W.A., Hogness D.S.;
RT "The E75 ecdysone-inducible gene responsible for the 75B early puff in
RT Drosophila encodes two new members of the steroid receptor
RT superfamily."
RL Genes Dev. 4:204-219(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sultion G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Champe M., Pfeiffer B.D.,
RA Wan R.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agayani A., An H.-O., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heitman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,
RA Jaitai M., Kalish F., Kappen G.H., Ke Z., Kemion J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtk R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

```

RA Wang Z.-Y., Maasartman D.A., Weinstock G.M., Weisenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of *Drosophila melanogaster*,"  
RL Science 287:2185-2195(2000).  
RN [3]  
RP GENOME REANNOTATION, AND ALTERNATIVE SPLICING.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminck J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Beitzencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.,  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review,"  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [4]  
RP FUNCTION, DEVELOPMENTAL STAGE, AND INDUCTION.  
RX MEDLINE=94038699; PubMed=8223281;  
RA Hueb F., Ruiz C., Richards G.,  
RT "Puffs and PCR: the in vivo dynamics of early gene expression during  
RT ecdysone responses in *Drosophila*,"  
RL Development 118:613-627(1993).  
CC -!- FUNCTION: Implicated in the regulation of ecdysone-triggered gene  
CC hierarchies. Probably plays a key role in mediating the regulation  
CC of the larval molt by 20-OH-ecdysone.  
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Name=A; Synonyms=E75B;  
CC IsoId=PI7672-1; Sequence=Displayed;  
CC Name=C; Synonyms=E75A;  
CC IsoId=PI7671-1; Sequence=External;  
CC Name=B; Synonyms=E75C;  
CC IsoId=PI3055-2; Sequence=External;  
CC Name=D;  
CC IsoId=PI7671-2; Sequence=External;  
CC Note=No experimental confirmation available;  
CC -!- DEVELOPMENTAL STAGE: In mid instar larvae salivary glands, levels  
CC are low during puff stage 1, increase during puff stages 2-4 and  
CC diminish from stage 5 onwards. In prepupae, isoform A is the  
CC predominant form during puff stage 19 and the transition to stage  
CC 20. By stage 3 it is present in the gut, Malpighian tubules and  
CC the fat body, levels persist beyond stage 11.  
CC -!- INDUCTION: The expression of this protein is developmentally  
CC regulated and is correlated with the 20-OH-ecdysone induced  
CC activity of puff 75B.  
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NRI  
CC subfamily.  
CC -!- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.  
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a  
CC frameshift in position 1320.  
CC  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC  
CC EMBL, X51549; CNA5924.1, ALT\_FRAME; mRNA.  
CC EMBL, AE003522; AN11688.1; -; Genomic DNA.  
CC PIR, B34598; B34598.  
CC HSSP, P20393; 1A6Y.  
CC SMR, P17672; 415-480.  
CC TRANSFAC, T01368; -;  
CC Flybase: FBgn0000568; E4p75B.  
CC GO, GO:0004879; F, ligand-dependent nuclear receptor activity; NAS.  
CC GO, GO:0019900; P, ecdysone (senen insecta); IMP.  
CC GO, GO:0035072; P, ecdysone-mediated induction of salivary gla. .; NAS.

DR GO:0007553; P, regulation of ecdysteroid metabolism; IMP.  
DR InterPro; IPR001628; Hrmr\_rcpt\_DNA\_bd.  
DR InterPro; IPR000536; Hrmr\_rcpt\_lig\_bd.  
DR InterPro; IPR000354; Involucrin\_rcp.  
DR InterPro; IPR003079; ROR\_receptor.  
DR InterPro; IPR001723; Stdhrm\_receptor.  
DR Pfam; PF00104; Hormone\_recep; 1.  
DR Pfam; PF00904; Involucrin; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PRO1293; RORNUCRCPT.  
DR PRINTS; PRO0398; STRDHOMER.  
DR PRINTS; PRO0047; STROIDFINGER.  
DR ProDom; PD000035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOL1; 1.  
DR SMART; SM00399; ZNF\_C4; 1.  
DR PROSITE; PS51030; NUCLEAR\_REC\_DBD\_2; 1;  
KW Alternative splicing; Developmental protein; DNA-binding;  
KW Metal-binding; Nuclear protein; Receptor; Transcription;  
KW Transcription regulation; Zinc; Zinc-finger.  
FT DNA\_BIND 384 474 Nuclear receptor.  
FT ZN\_FING 387 421 NR C4-type 1; degenerate.  
FT ZN\_FING 438 457 NR C4-type 2.  
FT COMPBIAS 6 324 Gln-rich.  
FT COMPBIAS 877 912 Gln-rich.  
FT COMPBIAS 1056 1259 Ser-rich.  
FT CONFLICT 206 206 S -> C (in Ref. 1).  
FT CONFLICT 244 244 L -> L (in Ref. 1).  
SO SEQUENCE 1355 AA; 147171 MW; 6EF19BACE56255 CRC64;  
  
QY 1 CXXXXXXC 9  
Db 1226 CXXXXXXC 1234  
  
RESULT 23  
Q9VU79 DROME PRELIMINARY; PRT; 1365 AA.  
AC Q9VU79;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE CG10231-PA.  
GN Name=Pdf1; ORFNames=CG10231;  
OS *Drosophila melanogaster* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; *Drosophila*.  
OX NCBI\_TaxId=7227;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benes P.V., Berman B.P., Bhattacharya D., Bolshakov S.,  
RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Crawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Dey S., Davis A.P., Dew I., Dietz S.M.,  
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.T., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Maassman D.A., Weinstock G.M., Weisenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
[2]  
RN NUCLEOTIDE SEQUENCE.  
RP MEDLINE=22426065; PubMed=12537568;  
RX Celinker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.,  
RA Svitek R., Taber P.B., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
RL melanogaster euchromatic genome sequence.";  
RN Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
[3]  
RN NUCLEOTIDE SEQUENCE.  
RP MEDLINE=22426070; PubMed=12537573;  
RX Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svitek R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celinker S.E.,  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
RL a genomic perspective.";  
RN Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).  
[4]  
RN NUCLEOTIDE SEQUENCE.  
RP MEDLINE=22426069; PubMed=12537572;  
RX Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tu Y.J., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bertencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.,  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RL systematic review.";  
RN Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
[5]  
RN NUCLEOTIDE SEQUENCE.  
RP Berkeley *Drosophila* Genome Project:  
RA Celinker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svitek R., Smith E.,  
RA Yu C., Rubin G.,  
RT "Drosophila melanogaster release 4 sequence.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[6]  
RN NUCLEOTIDE SEQUENCE.  
RP FlyBase;  
RG Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AEO03659; AAF53675.2; -; Genomic DNA.  
DR Ensembl: CG10231; Drosophila melanogaster.  
DR FlyBase: FBgn0032686; Pdel1.  
DR GO: GO:0004114; F:3',5'-cyclic-nucleotide phosphodiesterase a. . .; IEA.  
DR GO: GO:0003824; F:catalytic activity; IEA.  
DR GO: GO:0007165; P:signal transduction; IEA.

DR InterPro: IPR003018; GAF.  
DR InterPro: IPR003607; Met\_phos\_hydro.  
DR InterPro: IPR002073; PDEase.  
DR Pfam: PF01590; GAF; 2.  
DR Pfam: PF00233; PDEase\_I; 1.  
DR PRINTS: PR00387; PDESTERASE1.  
DR SMART: SM00065; GAF; 2.  
DR SMART: SM00471; HDC; 1.  
DR PROSITE: PS00126; PDEASE\_I; 1.  
SQ SEQUENCE 1365 AA, 151173 MW, 80562EFPD5012B56 CRC64;  
Query Match 72.0%; Score 18; DB 2; Length 1365;  
Best Local Similarity 22.2%; Fred. No. 6;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
DB 1138 CDSSTRASSC 1146  
QY 1 CXXXXXXXXX 9  
E75BB DROME STANDARD; PRT; 1412 AA.  
ID E75BB DROME STANDARD; PRT; 1412 AA.  
AC P13055; Q9VVM9;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 13-SEP-2005 (Rel. 48, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Ecdysone-induced protein 75B isoform B (E75-C).  
GN Name:E75B; Synonym=NR1D3; ORFNames=C68127;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RN NUCLEOTIDE SEQUENCE [MRNA], AND ALTERNATIVE SPLICING.  
RP STRAIN=Canton-S; TISSUE=Head;  
RX MEDLINE=90016778; PubMed=2508058;  
RA Feigl G., Gram M., Pongs O.,  
RT "A member of the steroid hormone receptor gene family is expressed in  
RL the 20-OH-ecdysone inducible puff 75B in *Drosophila melanogaster*.";  
RN Nucleic Acids Res. 17:7167-7178(1989).  
[2]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RP STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Boyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
RA Bailly R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Brottier P.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delber A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunker B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Folsler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,



Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 Abril J.F., Agayani A., An H.-J., Andrews-Pamkoc C., Baldwin D.,  
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borikova K., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Fabbis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Doug L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 Glodok A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,  
 Idali M., Kalish F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulip D., Lai Z.,  
 Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkulov G., Mishaeva N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Murthy D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of *Drosophila melanogaster*."; *Science* 287:2185-2195(2000).  
 [2]  
 NUCLEOTIDE SEQUENCE.  
 MEDLINE=22426065; PubMed=12537568;  
 Celinker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 George R.A., Hoskins R.A., Laverly T., Murthy D.M., Nelson C.R.,  
 Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 Weintraub G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
 melanogaster euchromatic genome sequence."; *Genome Biol.* 3:RESEARCH0079-RESEARCH0079(2002).  
 [3]  
 NUCLEOTIDE SEQUENCE.  
 MEDLINE=22426070; PubMed=12537573;  
 Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,  
 Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 Ashburner M., Celinker S.E.;  
 "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 a genomic perspective."; *Genome Biol.* 3:RESEARCH0084.1-RESEARCH0084.22(2002).  
 [4]  
 NUCLEOTIDE SEQUENCE.  
 MEDLINE=22426069; PubMed=12537572;  
 Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 Bettencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 Lewis S.E.;  
 "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 systematic review."; *Genome Biol.* 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 [5]  
 NUCLEOTIDE SEQUENCE.  
 Berkeley *Drosophila* Genome Project;  
 Kaminker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
 Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,  
 Yu C., Rubin G.;

RT "Drosophila melanogaster release 4 sequence."; *GenBank/DBJ* databases.  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [6]  
 NUCLEOTIDE SEQUENCE.  
 RP Flybase;  
 Submitter (MAR-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003436; AAF46107.1; -; Genomic DNA.  
 DR Ensembl; CG4790; *Drosophila melanogaster*.  
 DR Flybase; FBgn005390; CG4790.  
 DR Flybase; FBgn005390; fs(1)M3.  
 SO SEQUENCE 1793 AA; 202233 MW; DDC38A764B3001CD CRC64;  
 Query Match 72.0%; Score 18; DB 2; Length 1793;  
 Best Local Similarity 22.2%; Pred. No. 73;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 CXXXXXXC 9  
 DB 291 CTSSASASC 299  
 RESULT 26  
 Q8MP07\_DROME  
 ID Q8MP07; DROME PRELIMINARY; PRT; 1836 AA.  
 AC Q8MP07;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Poliole precursor.  
 GN Name=fs(1)M3; Synonyms=fs(1)ph;  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; *Drosophila*.  
 OC NCBI\_TaxId=7227;  
 RX MEDLINE=21956567; PubMed=11959840; DOI=10.1101/gad.223902;  
 RA Jimenez G., Gonzalez-Reyes A., Casanova J.;  
 RT "Cell surface proteins Nbsr1 and Poliole stabilize the Torso-like  
 extracellular determinant in *Drosophila* oogenesis."; *Cell*  
 RL Genes Dev. 16:913-918(2002).  
 DR EMBL; AJ487522; CAD31790.1; -; mRNA.  
 DR Ensembl; CG4790; *Drosophila melanogaster*.  
 DR Flybase; FBgn005390; CG4790.  
 DR Flybase; FBgn005390; fs(1)M3.  
 KW Signal.  
 FT SIGNAL 1 25 Potential.  
 FT CHAIN 26 1836 Poliole.  
 SQ SEQUENCE 1836 AA; 207449 MW; 4B0FC64DCBFS2EF CRC64;  
 Query Match 72.0%; Score 18; DB 2; Length 1836;  
 Best Local Similarity 22.2%; Pred. No. 74;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 CXXXXXXC 9  
 DB 291 CTSSASASC 299  
 RESULT 27  
 Q4Q515\_LEIMA  
 ID Q4Q515\_LEIMA PRELIMINARY; PRT; 1894 AA.  
 AC Q4Q515;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
 DE Calpain-like cysteine peptidase, putative (Cysteine peptidase, clan  
 DE ca, family C2, putative).  
 GN ORFNames=Lm1F32.0970;  
 GN Lm1F32.0970;  
 OS Leishmania major.  
 OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OC NCBI\_TaxId=5664;

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Fieldin;
RA Peacock C.S., Murphy L., Ivens A.C., Berrian M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrett B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CTO05269; CAJ08617.1; -; Genomic DNA.
SQ SEQUENCE 1894 AA; 199683 MW; 3F71DE2F53BF95BB CRC64;

Query Match 72.0%; Score 18; DB 2; Length 1894;
Best Local Similarity 22.2%; Pred. No. 75;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 9
DB 1785 CAAAAAACC 1793

RESULT 28
Q4S488.TETNG PRELIMINARY; PRT; 2087 AA.
ID Q4S488.TETNG
AC Q4S488;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Chromosome undetermined SCAF14743, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00024299001;
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetradodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallou O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costraz C., Bernot A.,
RA Michaud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castell V., Katinka N., Vacherie B.,
RA Biemont C., Spall S., Cattolico L., Poulain J., De Bernardis V.,
RA Cruaud C., Duprat S., Broctier P., Couranceau J.P., Gouzy J.,
RA Paria G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.,
RT "Genome duplication in the teleost fish Tetradodon nigroviridis reveals
RT the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG GenomeScope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
DR EMBL; CAAB01014743; CAG04544.1; -; Genomic DNA.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca_bd.
DR InterPro; IPR006209; EGF_1like.
DR InterPro; IPR000173; GAP_dhhdogenase.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR012680; Laminin_G_2.
DR Pfam; PF000028; Cadherin; 8.
DR Pfam; PF01049; Cadherin_C_1.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF02210; Laminin_G_2; 2.

```

```

DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 9.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00282; LamG; 2.
DR PROSITE; PS00232; CADHERIN_1; 6.
DR PROSITE; PS50268; CADHERIN_2; 9.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS50026; EGF_3; 2.
DR PROSITE; PS00071; GAPDH; UNKNOWN 1.
DR PROSITE; PS50025; LamG_DOMAIN; 2.
DR PROSITE; PS50025; LamG_DOMAIN; Transmembrane.
KW Calcium; EGF-like domain; Transmembrane.
FT NON TER 1
SQ SEQUENCE 2087 AA; 222663 MW; 2C5D0554ECDD7A1 CRC64;

Query Match 72.0%; Score 18; DB 2; Length 2087;
Best Local Similarity 22.2%; Pred. No. 78;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 9
DB 442 CSSSSAACC 450

RESULT 29
Q9L981.VIRCH PRELIMINARY; PRT; 50 AA.
ID Q9L981.VIRCH
AC Q9L981;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORF31.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=569b;
RX MEDLINE=20476433; PubMed=11021935;
RA Clark C.A., Putins L., Kaewrakon P., Focareta T., Manning P.A.;
RL "The Vibrio cholerae O1 chromosomal integrin."
RT Microbiology 146:2605-2612(2000).
DR EMBL; AF179596; AAF71194.1; -; Genomic DNA.
SQ SEQUENCE 50 AA; 5564 MW; 1418FC4DD3634E57 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 50;
Best Local Similarity 22.2%; Pred. No. 57;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 9
DB 28 CASLSSSC 36

RESULT 30
Q4RA29.TETNG PRELIMINARY; PRT; 71 AA.
ID Q4RA29.TETNG
AC Q4RA29;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Chromosome undetermined SCAF24619, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00037393001;
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetradodon.
OX NCBI_TaxID=99883;
RN [1]

```



RP NUCLEOTIDE SEQUENCE.  
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,  
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
RA Nicud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segreus B.,  
RA Dastilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
RA Anthouard V., Jubin C., Castell V., Katinka M., Vacherie B.,  
RA Blemont C., Skalli Z., Botticco L., Poulain J., De Bernardis V.,  
RA Cruaud C., Duprat G., Brotier P., Coutanceau J.P., Gouzy J.,  
RA Parra G., Lardier G., Chappie C., McKernan K.J., McSwan P., Bosak S.,  
RA Kellis M., Volt J.N., Guigo R., Zody M.C., Mesirov J.,  
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
RA Lauder V., Schachter V., Queller F., Saurin W., Scarpelli C.,  
RA Winkler P., Lander E.S., Weissenbach J., Roest Crolius H.;  
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
RT the early vertebrate proto-karyotype".  
RL Nature 431:946-957(2004).  
RN (2)  
RP NUCLEOTIDE SEQUENCE.  
RG Genoscope; Whitehead Institute Centre for Genome Research;  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; CAEB01024619; CAG14754.1; -; Genomic\_DNA.  
FT NON TER 1 1  
SQ SEQUENCE 71 AA; 8036 MW; 736AF2E349B9AC50 CRC64;  
QY 1 CXXXXXXC 9  
Db 27 CSTRSPARC 35

RESULT 31  
Q8VD96 MESAU PRELIMINARY; PRT; 96 AA.  
ID Q8VD96;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE 10 kDa protein precursor.  
GN Name=CC10;  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridae; Cricetinae; Mesocricetus.  
OX NCBI\_TaxID=10036;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Lung;  
RA Gutierrez-Sagat R., Nieto A.;  
RT "Cloning and sequencing of a cDNA encoding hamster uteroglobin/clara  
RT cell 10 kDa protein.";  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L37041; AAL31349.1; -; Genomic\_DNA.  
DR HSSP; P17559; 1UTR.  
DR SMR; Q8VD96; 20-96.  
DR GO; GO:0005496; F:steroid binding; IEA.  
DR InterPro; IPR006039; Utergl.  
DR InterPro; IPR003628; Uteroglbn\_sub.  
DR InterPro; IPR003329; Uteroglbn\_subf.  
DR InterPro; IPR006038; Uteroglbn\_subf.  
DR PANTHER; PTHR10136; Uteroglbn\_sub; 1.  
DR Pfam; PF01039; Uteroglbn; 1.  
DR PRINTS; PR00486; UTEROGLBIN.  
DR PRODOM; PD012475; Uteroglbn\_sub; 1.  
DR SMART; SM00096; UTG; 1.  
DR DR PROSITE; PS00404; UTEROGLBIN\_2; 1.  
KW Signal.  
FT SIGNAL 1 18 Potential.

SQ SEQUENCE 96 AA; 10509 MW; 5EB9CBDD46143389 CRC64;  
QY 1 CXXXXXXC 9  
Db 16 CSSASDTC 24

RESULT 32  
Q4P4H6 USTWA  
ID Q4P4H6 USTWA PRELIMINARY; PRT; 108 AA.  
AC Q4P4H6;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=UM04987.1;  
OS Ustilago maydis 521.  
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;  
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.  
OX NCBI\_TaxID=237631;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=521;  
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,  
RA Altshuler M., Allen N., Allen T., An P., Anderson M., Anderson S.,  
RA Arachchi H., Amburster J., Bachantansang P., Baldwin W., Barry A.,  
RA Bayul T., Blitshstein B., Bloom T., Bye J., Boguslavsky L.,  
RA Borowsky M., Boukhalter B., Brunache A., Butler J., Calixte N.,  
RA Calvo S., Camarata J., Campo K., Chang J., Chesnatsang Y., Citroen M.,  
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,  
RA David R., Dawes T., Degray S., Dodge S., Dooley K., Dorje P.,  
RA Dorjee K., Dorris L., Dufey N., Dupes A., Elkins T., Engels R.,  
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,  
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gherre S.,  
RA Glatzer A., Govette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,  
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,  
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,  
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamywasalis M., Karlsson E.,  
RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutis K.,  
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,  
RA Lindblad-Toh K., Liu X., Lokitsang T., Lokitsang Y., Lucien O.,  
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,  
RA Manning J., Marabella R., Marr K., Matthews C., Mauceli E.,  
RA McCarthy M., McDonough S., McGhee T., Melidrim J., Meneu L.,  
RA Mesirov J., Mihalov A., Mihova T., Mikkelson T., Mienga V., Moru K.,  
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,  
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,  
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pigani B.,  
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,  
RA Rector A., Richardson S., Rice C., Rodriguez J., Rogers J., Rogov P.,  
RA Ruttman M., Schupbach R., Seaman C., Settippalli S., Sharpe T.,  
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sounez C.,  
RA Spencer B., Stalker J., Strange-Thomann N., Stavropoulos S.,  
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuina P.,  
RA Tenzing P., Teftaye S., Theodore J., Thoutang Y., Topham K.,  
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,  
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,  
RA Wengli T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,  
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembek L.,  
RA Zimmer A., Zody M., Lander E.;  
RT "The genome sequence of Ustilago maydis".  
RT Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AACP01000184; EAK85760.1; -; Genomic\_DNA.  
DR DR Hypothetical protein.  
KW Signal.  
SQ SEQUENCE 108 AA; 11407 MW; E641B76D8BC16E7A CRC64;



Query Match 68.0%; Score 17; DB 2; Length 108;  
 Best Local Similarity 22.2%; Pred. No. 78;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 DB 14 CTSANATC 22

## RESULT 33

Q9ZSP6\_ARATH PRELIMINARY; PRT; 108 AA.

AC Q9ZSP6;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 27, Last annotation update)  
 DE Putative cell wall-plasma membrane disconnecting CLCT protein.  
 GN Name=ARL1B;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC Eurosid 1; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;

RP NUCLEOTIDE SEQUENCE.

RA Neutboom L.W., van der Zaal E.J.;  
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.

RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,  
 RA Nakajima M., Enju A., Kamiya A., Natusaka M., Carninci P., Kawai J.,  
 RA Hayashizaki Y., Shinozaki K.;  
 RL Submitted (Nov-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF098631; AAD12259.1; -; Genomic DNA.  
 DR EMBL; AK118720; BAC43314.1; -; mRNA.  
 DR HSSP; P24337; IHP.

DR InterPro; IPR003612; AAI.

DR Pfam; PF00234; Tryp\_alpha\_amy1; 1.

DR SMART; SM00499; AAI; 1.1194 MW; 8CD9F6CAE2577D2 CRC64;

QY 1 CXXXXXXC 9  
 DB 20 CTSATGTC 28

Query Match 68.0%; Score 17; DB 2; Length 108;  
 Best Local Similarity 22.2%; Pred. No. 78;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

## RESULT 34

Q72D14\_DESVH PRELIMINARY; PRT; 115 AA.

AC Q72D14;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein.

GN OrderedlocusNames=DVU1117;  
 OS Desulfotribrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303).

OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfotribionales;  
 OC Desulfotribriaceae; Desulfotribrio.

OC NCBI\_TaxID=882;

RP NUCLEOTIDE SEQUENCE.

RA PubMed=15077118; DOI=10.1038/nbt959;

RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,  
 RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,  
 RA Daugherty S.C., Deboy R.T., Dodson R.J., Durkin A.S., Madupu R.,  
 RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,  
 RA Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,

RA Dimirov G., Hance M., Tran K., Khouri H.M., Gill J., Uterback T.R.,  
 RA Feldlyum T.V., Wall J.D., Voordouw G., Fraser C.M.;  
 RT "The genome sequence of the anaerobic, sulfate-reducing bacterium  
 RT Desulfotribrio vulgaris Hildenborough.";

RL Nat. Biotechnol. 22:554-559(2004).

DR EMBL; AE017313; AAS95597.1; -; Genomic DNA.  
 DR TIGR; DVU1117; -;  
 KM Complete proteome; Hypothetical protein.

QY 1 CXXXXXXC 9  
 DB 70 CDTATASC 78

Query Match 68.0%; Score 17; DB 2; Length 115;  
 Best Local Similarity 22.2%; Pred. No. 80;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 DB 70 CDTATASC 78

## RESULT 35

Q6VZRI\_CNPV PRELIMINARY; PRT; 117 AA.

AC Q6VZRI;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE CNP0086 TNFR-like protein.

GN Name=CNP0086;

OS Canaripox virus (CNPV).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Avipoxvirus.

OC NCBI\_TaxID=44088;

RP NUCLEOTIDE SEQUENCE.

RA STRAIN=ATCC VR-111;  
 RC PubMed=14671117; DOI=10.1128/JVI.78.1.353-366.2004;  
 RA Tulman E.R., Alonso C.L., Lu Z., Zsak J., Kutish G.F., Rock D.L.;  
 RT "The genome of canaripox virus.";

RL J. Virol. 78:353-366(2004).

DR EMBL; AY318871; AAR83432.1; -; Genomic DNA.  
 GO; GO:0004872; Fireceptor activity; IEA.

DR InterPro; IPR006209; EGF-like.

DR Pfam; PF00020; TNFR\_c6; 1.

DR SMART; SM00208; TNFR; 2.

DR PROSITE; PS01186; EGF\_2; UNKNOWN\_1.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE; PS00500; TNFR\_NGFR\_2; 1.

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

Query Match 68.0%; Score 17; DB 2; Length 117;  
 Best Local Similarity 22.2%; Pred. No. 80;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

QY 1 CXXXXXXC 9  
 DB 50 CTSSTTTC 58

```

RC STRAIN=Portalegre 90;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
  fever virus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028332; AAY46118.1; -; Genomic_DNA.
FT NON_TER 1 1
FT SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 22.2%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 37
Q4TWL3_ASF PRELIMINARY; PRT; 119 AA.
ID Q4TWL3;
AC Q4TWL3;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
ON NCBI_TaxID=10497;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Lisbon 60;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
  fever virus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028331; AAY46117.1; -; Genomic_DNA.
FT NON_TER 1 1
FT SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 22.2%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 38
Q4TWL5_ASF PRELIMINARY; PRT; 119 AA.
ID Q4TWL5;
AC Q4TWL5;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
ON NCBI_TaxID=10497;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Coimbra 87;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
  fever virus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028329; AAY46115.1; -; Genomic_DNA.
FT NON_TER 1 1
FT SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

```

```

FT NON_TER 119 119
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 22.2%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 39
Q4TWL6_ASF PRELIMINARY; PRT; 119 AA.
ID Q4TWL6;
AC Q4TWL6;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
ON NCBI_TaxID=10497;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Mafra 86;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
  fever virus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028328; AAY46114.1; -; Genomic_DNA.
FT NON_TER 1 1
FT SEQUENCE 119 AA; 11863 MW; 3A062A08CA8D31F CRC64;

Query Match
Best Local Similarity 22.2%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 9
DB 33 CASTGASTC 41

RESULT 40
Q4TWL7_ASF PRELIMINARY; PRT; 119 AA.
ID Q4TWL7;
AC Q4TWL7;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE 9RL protein (Fragment).
GN Name=B602L;
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
ON NCBI_TaxID=10497;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Barrancos 93;
RA Duarte M.M., Cruz M.B.;
RT "B602L (9RL) partial gene of Portuguese isolates of African swine
  fever virus.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ028327; AAY46113.1; -; Genomic_DNA.
FT NON_TER 1 1
FT SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match
Best Local Similarity 22.2%; Score 17; DB 2; Length 119;
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 1 CXXXXXXC 9  
DB 33 CASTGASTC 41

## RESULT 41

ID Q4TWL8 ASF PRELIMINARY; PRT; 119 AA.  
AC Q4TWL8;  
DT 13-SEP-2005 (T-EMBLrel. 31, Created)  
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)  
DE 9RL protein (Fragment).  
GN Name=B602L;  
OS African swine fever virus (ASFV).  
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.  
OX NCBI\_TaxID=10497;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Almodovar 99;  
RA Duarte M.M., Cruz M.B.;  
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; DQ028326; AAY46112.1; -; Genomic\_DNA.  
FT NON TER 1  
FT NON TER 119  
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 119;  
Best Local Similarity 22.2%; Pred. No. 81;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 33 CASTGASTC 41

## RESULT 42

ID Q4TWL9 ASF PRELIMINARY; PRT; 119 AA.  
AC Q4TWL9;  
DT 13-SEP-2005 (T-EMBLrel. 31, Created)  
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)  
DE 9RL protein (Fragment).  
GN Name=B602L;  
OS African swine fever virus (ASFV).  
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.  
OX NCBI\_TaxID=10497;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Almodovar 99/NE1;  
RA Duarte M.M., Cruz M.B.;  
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; DQ028325; AAY46111.1; -; Genomic\_DNA.  
FT NON TER 1  
FT NON TER 119  
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 119;  
Best Local Similarity 22.2%; Pred. No. 81;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 33 CASTGASTC 41

RESULT 43  
Q4TWM0 ASF

ID Q4TWM0 ASF PRELIMINARY; PRT; 119 AA.

AC Q4TWM0;  
DT 13-SEP-2005 (T-EMBLrel. 31, Created)  
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)  
DE 9RL protein (Fragment).  
GN Name=B602L;  
OS African swine fever virus (ASFV).  
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.  
OX NCBI\_TaxID=10497;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Almodovar 99/E2;  
RA Duarte M.M., Cruz M.B.;  
RT "B602L (9RL) partial gene of Portuguese isolates of African swine fever virus."  
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; DQ028324; AAY46110.1; -; Genomic\_DNA.  
FT NON TER 1  
FT NON TER 119  
SQ SEQUENCE 119 AA; 11864 MW; DDC8B817463AC420 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 119;  
Best Local Similarity 22.2%; Pred. No. 81;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 33 CASTGASTC 41

## RESULT 44

ID O5YPNS NOCFA PRELIMINARY; PRT; 125 AA.  
AC O5YPNS;  
DT 25-OCT-2004 (T-EMBLrel. 28, Created)  
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)  
DE 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocustNames=nfs50040;  
OS Nocardia farcinica.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Nocardiaceae; Nocardia.  
OX NCBI\_TaxID=37329;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=IFM 10152;  
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;  
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K., Shiba T., Hattori M.;  
RT "The complete genomic sequence of Nocardia farcinica IFM 10152."  
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).  
DR EMBL; AP006618; BAD59856.1; -; Genomic\_DNA.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 125 AA; 12749 MW; 1FD69A1CADF924EB CRC64;

Query Match 68.0%; Score 17; DB 2; Length 125;  
Best Local Similarity 22.2%; Pred. No. 83;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CXXXXXXC 9  
DB 96 CRSSSATTC 104

## RESULT 45

ID Q4T2U4 TETNG PRELIMINARY; PRT; 133 AA.  
AC Q4T2U4;  
DT 13-SEP-2005 (T-EMBLrel. 31, Created)  
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)  
DE 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)  
DE Chromosome 12 SCAFI0174, whole genome shotgun sequence.

DE (Fragment).  
 GN ORNames=GSFENG00008171001;  
 OC Tetraodon nigroviridis (Green puffer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Tetraodon.  
 OC NCBI\_TaxID=99883;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,  
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
 RA Nicard S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
 RA Anthouard V., Jubin C., Castelli V., Karinka M., Vacherie B.,  
 RA Biemont C., Skalli Z., Catolico L., Poulin J., De Berardinis V.,  
 RA Parra G., Lardier G., Chaptier C., Coutanceau J.P., Gouzy J.,  
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
 RA Lauder V., Schacherer V., Quetier F., Saurin W., Scarpelli C.,  
 RA Wincker P., Lander E.S., Weissbach U., Roest Crollius H.,  
 RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
 RT the early vertebrate proto-karyotype."  
 RL Nature 431:946-957(2004).  
 RN [2]

RP NUCLEOTIDE SEQUENCE.  
 RG Genoscope; Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; CAAB0101074; CAF92788.1; -; Genomic\_DNA.  
 FT NON\_FER 1  
 FT SEQUENCE 133 AA; 14711 MW; 528890284B8084F4 CRC64;

Query Match 68.0%; Score 17; DB 2; Length 133;  
 Best local similarity 22.2%; Pred. No. 85;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 9  
 Db 122 CQATSSPTC 130

RESULT 46  
 Q6ZRF5\_HUMAN  
 ID Q6ZRF5\_HUMAN PRELIMINARY; PRT; 134 AA.  
 AC Q6ZRF5;  
 DT 05-JUN-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUN-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein FLJ46388.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;  
 OC Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TSSUE=Flymus;  
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,  
 RA Fukunumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,  
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,  
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,  
 RA Yagetsuna M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,  
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahara K.,  
 RA Masuhira Y., Nagai K., Isogai T.,  
 RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK128253; BAC87355.1; -, mRNA.  
 SQ SEQUENCE 134 AA; 14481 MW; 03BA7E349C87BC6A CRC64;

Query Match 68.0%; Score 17; DB 2; Length 134;  
 Best local similarity 22.2%; Pred. No. 85;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 9  
 Db 27 CSASAAQC 35

RESULT 47  
 Q9CWC3\_MOUSE  
 ID Q9CWC3\_MOUSE PRELIMINARY; PRT; 147 AA.  
 AC Q9CWC3;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Mus musculus ES cells cDNA, RIKEN full-length enriched library,  
 DE clone:2410101B22 product:LENGTHS PROTEIN homolog.  
 GN Name=Length5;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RT Carninci P., Hayashizaki Y.,  
 RT "High-efficiency full-length cDNA cloning."  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J;  
 RX MEDLINE=21085650; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai T., Shingawa A., Shidara K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Komoto H., Adachi U., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,  
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsumoto H. A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glisic C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirml L. M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,  
 RA Brownstein M. J., Bulc C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinich S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mondaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyokawa K., Wang K. H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [3]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J;  
 RA The PANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team,  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 RN [4]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J;  
 RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
 RT "Normalization and subtractions of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes."  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CS7BL/6J;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagoka S., Sasaki N., Carninci P.,  
 RA Kuno H., Akiyama J., Nishi K., Katsunari T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kaishwagi K.,  
 RA Fujiwaki S., Inoue K., Togawa K., Izawa M., Ohara E., Matsumi M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Ozaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CS7BL/6J;  
 RA Aachai J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Aakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numata R., Ono M.,  
 RA Ohtsaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK010759; BAB27164.1; -, mRNA.  
 DR Ensembl; ENSMUSG0000035585; Mus musculus.  
 DR MGI; MGI.1913328; Lenz5.  
 SO SEQUENCE 147 AA; 15856 MW; 988434686F9EA74 CRC64;  
 Query Match 68.0%; Score 17; DB 2; Length 147;  
 Best Local Similarity 22.2%; Pred. No. 88;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 Oy 1 CXXXXXXC 9  
 Db 19 CATASTCTC 27

RESULT 48  
 Q6F340 ORYSA PRELIMINARY; PRT; 155 AA.  
 ID Q6F340 ORYSA PRELIMINARY; PRT; 155 AA.  
 AC Q6F340;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DE Hypothesized protein OSUNBa0088106.14.  
 GN Name=OSUNBa0088106.14;  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OC NCBI\_TaxID=39947;  
 OX [1]  
 RN NUCLEOTIDE SEQUENCE.  
 RP Chow T.-S., Hsiang Y.-I. C., Chen C.-S., Chen H.-H., Liu S.-M.,  
 RA Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,  
 RA Chao Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,  
 RA Huang J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Lue H.-L.,  
 RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,  
 RA Wu H.-P., Shaw J.-F.;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC129718; AAT69648.1; -, Genomic DNA.  
 DR Gramene; Q6F340; -  
 KM Hypothesized protein.  
 SO SEQUENCE 155 AA; 16979 MW; 0EADP4C906C39295 CRC64;  
 Query Match 68.0%; Score 17; DB 2; Length 155;  
 Best Local Similarity 22.2%; Pred. No. 90;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 1 CXXXXXXC 9  
 Db 146 CRASASTC 154

RESULT 49  
 Q4TWL4 ASF PRELIMINARY; PRT; 145 AA.  
 ID Q4TWL4 ASF PRELIMINARY; PRT; 145 AA.  
 AC Q4TWL4;  
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
 DE 9RL protein (Fragment).  
 GN Name=B602L;  
 OS African swine fever virus (ASFV).  
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfarivirus.  
 OC NCBI\_TaxID=10497;  
 OX [1]  
 RN NUCLEOTIDE SEQUENCE.  
 RC STRAIN=NH/P68;  
 RP Duarte M.M., Cruz M.B.;  
 RA "B602L (9RL) partial gene of Portuguese isolates of African swine  
 RT fever virus.";  
 RT Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; DQ028330; AAY46116.1; -, Genomic DNA.  
 DR NON\_TER 1 155  
 FT NON\_TER 1 155  
 SO SEQUENCE 155 AA; 15125 MW; 31AD6598783CE06 CRC64;  
 Query Match 68.0%; Score 17; DB 2; Length 155;  
 Best Local Similarity 22.2%; Pred. No. 90;  
 Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 Oy 1 CXXXXXXC 9  
 Db 69 CASTGASTC 77

RESULT 50  
 C1025 HUMAN STANDARD; PRT; 157 AA.  
 ID C1025 HUMAN STANDARD; PRT; 157 AA.  
 AC Q81W50; Q5T591; Q8T428;  
 DT 13-SEP-2005 (Rel. 48, Created)  
 DT 13-SEP-2005 (Rel. 48, Last sequence update)  
 DT 13-SEP-2005 (Rel. 48, Last annotation update)  
 DE Protein C9orf25.  
 GN Name=C9orf25;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;  
 OC Homo.  
 OC NCBI\_TaxID=9606;  
 OX [1]  
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RP PubMed=15164053; DOI=10.1038/nature02465;  
 RX Hummray S.J., Oliver K., Hunt A.R., Plumb R.W., Loveland J.B.,  
 RA Hummray S.J., Oliver K., Hunt A.R., Plumb R.W., Loveland J.B.,  
 RA Howe K.L., Andrews T.D., Searle S., Hunt S.E., Scott C.E., Jones M.C.,  
 RA Ainscough R., Almeida J.P., Ambrose K.D., Ashwell R.I.S.,  
 RA Babiker A.K., Babage S., Bagunley C.L., Bailey J., Banerjee R.,  
 RA Barker D.J., Barlow K.F., Bates K., Beasley H., Beasley O., Bird C.P.,  
 RA Bray-Allen S., Brown A.J., Brown J.Y., Burford D., Burrill W.,  
 RA Burton J., Carter C., Carter N.P., Chapman J.C., Chen Y., Clarke G.,  
 RA Clark S.Y., Clegg C.M., Clegg S., Collier R.E., Corby N., Croxier M.,  
 RA Cummings A.T., Davies J., Dhami P., Dunn M., Dutta I., Dyer L.W.,  
 RA Earlrow M.E., Faulkner L., Fleming C.J., Franklin A.,  
 RA Frankland J.A., French L., Fricker D.G., Garner P., Garnett J.,  
 RA Ghori J., Gilbert J.G.R., Glison C., Grafham D.V., Gribble S.,  
 RA Griffiths C., Griffiths-Jones S., Grocock R., Guy J., Hall R.E.,  
 RA Hammond S., Harley J.L., Harrison E.S.I., Hart E.A., Heath P.D.,  
 RA Henderson C.D., Hopkins B.L., Howard P.J., Howden P.J., Huckle E.,  
 RA Johnson C., Johnson D., Joy A.A., Kay M., Keenan S., Kershaw J.K.,  
 RA Kimberley A.M., King A., Knights A., Laird G.K., Langford C.,

RA Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C., Lloyd D.M.,  
RA Lovell J., Martin S., Mashreghi-Mohammadi M., Matthews L., McLaren S.,  
RA McElay K.E., McMurray A., Milne S., Nickerson T., Nisbett J.,  
RA Nordiek G., Pearce A.V., Peck A.I., Porter K.M., Pandian R.,  
RA Pelan S., Phillimore B., Povey S., Ramsey Y., Rand V., Scharfe M.,  
RA Sehra H.K., Showkeen R., Sims S.K., Skuce C.D., Smith M.,  
RA Steward C.A., Swarbreck D., Sycamore N., Tester J., Thorpe A.,  
RA Tracey A., Tromans A., Thomas D.W., Wall M., Wallis J.M., West A.P.,  
RA Whitehead S.L., Willey D.L., Williams S.A., Wilming L., Wray P.W.,  
RA Young L., Aehuet J.L., Cousion A., Blocker H., Durbin R.,  
RA Sulston J.E., Hubbard T., Jackson M.J., Bentley D.R., Beck S.,  
RA Rogers J., Dunham I.,  
RT "DNA sequence and analysis of human chromosome 9.";  
RL Nature 429:369-374(2004).  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,  
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uebin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL; AL160270; CA113153.1; ALT\_INIT; Genomic DNA.  
DR EMBL; AL156494; CA113153.1; JOINED; Genomic DNA.  
DR EMBL; AL156494; CA115974.1; ALT\_INIT; Genomic DNA.  
DR EMBL; AL160270; CA115974.1; JOINED; Genomic DNA.  
DR EMBL; BC025348; AAH25348.1; -; mRNA.  
DR EMBL; BC041009; AAH41009.1; -; mRNA.  
DR EMBL; ENSG00000164970; Homo sapiens.  
DR HNCI; HGNC:19920; C9orf25.  
FT CONFLICT 106 106 Missing (in Ref. 2; AAH41009).  
SQ SEQUENCE 157 AA; 17365 MW; 577AFAA46B7DC74 CRC64;  
  
Query Match 68.0%; Score 17; DB 1; Length 157;  
Best Local Similarity 22.2%; Pred. No. 90;  
Matches 2; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXC 9  
DB 146 COATSTRAC 154

Search completed: January 4, 2006, 16:09:53  
Job time : 59.7739 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 14:57:44 ; Search time 68.4348 Seconds  
(without alignments)  
64.204 Million cell updates/sec

Title: US-09-932-322-11

Perfect score: 26

Sequence: 1 CXXXXXXXXX 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 244163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 244163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

1: Geneseq\_21: \*  
2: geneseqp1980s: \*  
3: geneseqp1980s: \*  
4: geneseqp2000s: \*  
5: geneseqp2000s: \*  
6: geneseqp2003as: \*  
7: geneseqp2003as: \*  
8: geneseqp2004s: \*  
9: geneseqp2005s: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	69.2	30	9	AD247471
2	18	69.2	89	8	ADP30533
3	18	69.2	89	8	ADP30531
4	18	69.2	109	4	AAH93527
5	18	69.2	109	8	ADL31230
6	18	69.2	109	8	ADP30690
7	18	69.2	144	8	ADP31474
8	18	69.2	167	7	ABO69151
9	18	69.2	169	8	ADP312185
10	18	69.2	171	8	ADP30794
11	18	69.2	171	8	ADP30793
12	18	69.2	183	8	ADP30806
13	18	69.2	191	7	ABO76619
14	18	69.2	214	9	ADZ56213
15	18	69.2	228	8	ADP30921
16	18	69.2	252	8	ADP31485
17	18	69.2	258	8	ADP30479
18	18	69.2	264	8	ADP31412
19	18	69.2	270	8	ADP30500
20	18	69.2	270	8	ADP31321
21	18	69.2	294	8	ADP31473
22	18	69.2	297	8	ADP31192
23	18	69.2	328	4	AAH25285
24	18	69.2	339	8	ADP30702

25	18	69.2	348	8	ADP31441	Adp31441 Human sec
26	18	69.2	356	4	ABG21039	Abg21039 Novel hum
27	18	69.2	357	8	ADP31267	Adp31267 Human sec
28	18	69.2	357	8	ADP30505	Adp30505 Human sec
29	18	69.2	360	8	ADP31439	Adp31439 Human sec
30	18	69.2	390	8	ADP31218	Adp31218 Human sec
31	18	69.2	421	8	ADP31159	Adp31159 Human sec
32	18	69.2	426	8	ADP31495	Adp31495 Human sec
33	18	69.2	438	7	ABO82561	AbO82561 Pseudomon
34	18	69.2	453	8	ADP31465	Adp31465 Human sec
35	18	69.2	471	8	ADP30854	Adp30854 Human sec
36	18	69.2	525	8	ADP31227	Adp31227 Human sec
37	18	69.2	549	8	ADP30855	Adp30855 Human sec
38	18	69.2	555	8	ADP31416	Adp31416 Human sec
39	18	69.2	558	8	ADP31417	Adp31417 Human sec
40	18	69.2	588	8	ADP31699	Adp31699 Human sec
41	18	69.2	591	9	ADZ56209	Adz56209 Human KIA
42	18	69.2	591	9	ADZ56223	Adz56223 Human KIA
43	18	69.2	604	8	ADP30865	Adp30865 Human sec
44	18	69.2	604	8	ADP30940	Adp30940 Human sec
45	18	69.2	604	8	ADP30941	Adp30941 Human sec
46	18	69.2	605	8	ADP30507	Adp30507 Human sec
47	18	69.2	617	8	ADP31657	Adp31657 Human sec
48	18	69.2	626	6	ABR41586	AbR41586 Human DIR
49	18	69.2	645	8	ADP30858	Adp30858 Human sec
50	18	69.2	653	3	AAy94907	Aay94907 Human sec
51	18	69.2	653	5	ADZ73254	Adz73254 Human KIA
52	18	69.2	653	5	ADW42108	Adw42108 Protein e
53	18	69.2	653	7	ADB69077	AdB69077 TAT290. 1
54	18	69.2	653	8	ADZ510948	Adz510948 Human KIA
55	18	69.2	653	9	ADZ56204	Adz56204 Human KIA
56	18	69.2	653	9	ADZ56214	Adz56214 Human KIA
57	18	69.2	660	8	ADP31677	Adp31677 Human sec
58	18	69.2	665	8	ADP30571	Adp30571 Human sec
59	18	69.2	681	8	ADP31582	Adp31582 Human sec
60	18	69.2	690	8	ADP30780	Adp30780 Human sec
61	18	69.2	708	8	ADP31191	Adp31191 Human sec
62	18	69.2	711	8	ADP31440	Adp31440 Human sec
63	18	69.2	728	8	ADP30508	Adp30508 Human sec
64	18	69.2	739	8	ADP31196	Adp31196 Human sec
65	18	69.2	750	8	ADP30576	Adp30576 Human sec
66	18	69.2	757	8	ADP30925	Adp30925 Human sec
67	18	69.2	762	8	ADP30891	Adp30891 Human sec
68	18	69.2	771	8	ADP31469	Adp31469 Human sec
69	18	69.2	772	8	ADP30936	Adp30936 Human sec
70	18	69.2	774	8	ADP30506	Adp30506 Human sec
71	18	69.2	774	8	ADP30780	Adp30780 Human sec
72	18	69.2	774	8	ADP30781	Adp30781 Human sec
73	18	69.2	795	8	ADP31352	Adp31352 Human sec
74	18	69.2	804	8	ADP31291	Adp31291 Human sec
75	18	69.2	813	8	ADP31282	Adp31282 Human sec
76	18	69.2	813	8	ADP30649	Adp30649 Human sec
77	18	69.2	834	8	ADP30738	Adp30738 Human sec
78	18	69.2	852	8	ADP30664	Adp30664 Human sec
79	18	69.2	856	8	ADP30972	Adp30972 Human sec
80	18	69.2	865	8	ADP30566	Adp30566 Human sec
81	18	69.2	882	8	ADP31688	Adp31688 Human sec
82	18	69.2	888	8	ADP30971	Adp30971 Human sec
83	18	69.2	892	8	ADP31578	Adp31578 Human sec
84	18	69.2	930	8	ADP31444	Adp31444 Human sec
85	18	69.2	933	8	ADP31486	Adp31486 Human sec
86	18	69.2	947	8	ADP30937	Adp30937 Human sec
87	18	69.2	960	8	ADP31471	Adp31471 Human sec
88	18	69.2	960	8	ADP31470	Adp31470 Human sec
89	18	69.2	1017	6	AAO26719	AAO26719 SR protei
90	18	69.2	1033	8	ADP30984	Adp30984 Human sec
91	18	69.2	1065	8	ADP30661	Adp30661 Human sec
92	18	69.2	1065	8	ADP30973	Adp30973 Human sec
93	18	69.2	1065	8	ADP31482	Adp31482 Human sec
94	18	69.2	1090	8	ADP310187	Adp310187 Human pol
95	18	69.2	1092	8	ADP31358	Adp31358 Human sec
96	18	69.2	1110	8	ADP31430	Adp31430 Human sec
97	18	69.2	1116	8	ADP31692	Adp31692 Human sec

98	18	69.2	1132	8	ADP30952	Adp30952	Human	sec	171	18	69.2	3339	8	ADP31219	Adp31219	Human	sec
99	18	69.2	1133	8	ADP30931	Adp30931	Human	sec	172	18	69.2	3390	8	ADP31148	Adp31148	Human	sec
100	18	69.2	1133	8	ADP30928	Adp30928	Human	sec	173	18	69.2	3411	8	ADP30667	Adp30667	Human	sec
101	18	69.2	1133	8	ADP30933	Adp30933	Human	sec	174	18	69.2	3447	8	ADP31112	Adp31112	Human	sec
102	18	69.2	1133	8	ADP30939	Adp30939	Human	sec	175	18	69.2	3477	8	ADP30704	Adp30704	Human	sec
103	18	69.2	1142	8	ADP30929	Adp30929	Human	sec	176	18	69.2	3440	6	ABU88256	Abu88256	Novel	hum
104	18	69.2	1165	8	ADP30879	Adp30879	Human	sec	177	18	69.2	4440	6	ABU90135	Abu90135	Novel	hum
105	18	69.2	1173	8	ADP31155	Adp31155	Human	sec	178	18	69.2	4440	6	ABU96437	Abu96437	Novel	hum
106	18	69.2	1173	8	ADP31034	Adp31034	Human	sec	179	18	69.2	4440	6	ABU99046	Abu99046	Novel	hum
107	18	69.2	1209	8	ADP30782	Adp30782	Human	sec	180	18	69.2	4440	6	ABU98261	Abu98261	Novel	hum
108	18	69.2	1222	8	ADP30501	Adp30501	Human	sec	181	18	69.2	4440	6	ABU91967	Abu91967	Novel	hum
109	18	69.2	1233	8	ADP30523	Adp30523	Human	sec	182	18	69.2	4440	6	ABU85271	Abu85271	Novel	hum
110	18	69.2	1239	8	ADP31297	Adp31297	Human	sec	183	18	69.2	4440	6	ABU00410	Abu00410	Novel	hum
111	18	69.2	1305	8	ADP31389	Adp31389	Human	sec	184	18	69.2	4440	6	ABU88961	Abu88961	Novel	hum
112	18	69.2	1358	8	ADP30995	Adp30995	Human	sec	185	18	69.2	4440	6	ABU06457	Abu06457	Novel	hum
113	18	69.2	1365	8	ADP31035	Adp31035	Human	sec	186	18	69.2	4440	6	ABU95517	Abu95517	Novel	hum
114	18	69.2	1371	8	ADP30876	Adp30876	Human	sec	187	18	69.2	4440	6	ABU95207	Abu95207	Novel	hum
115	18	69.2	1387	8	ADP30946	Adp30946	Human	sec	188	18	69.2	4440	6	ABU90755	Abu90755	Novel	hum
116	18	69.2	1431	8	ADP31609	Adp31609	Human	sec	189	18	69.2	4440	6	ABU93917	Abu93917	Novel	hum
117	18	69.2	1440	8	ADG34533	Adg34533	GlucoCort		190	18	69.2	4440	6	ABU86191	Abu86191	Novel	hum
118	18	69.2	1472	8	ADP31611	Adp31611	Human	sec	191	18	69.2	4440	6	ABU82046	Abu82046	Novel	hum
119	18	69.2	1480	8	ADP30557	Adp30557	Human	sec	192	18	69.2	4440	6	ABU94227	Abu94227	Novel	hum
120	18	69.2	1518	8	ADP31532	Adp31532	Human	sec	193	18	69.2	4440	6	ABU00100	Abu00100	Novel	hum
121	18	69.2	1550	8	ADP30567	Adp30567	Human	sec	194	18	69.2	4440	6	ABU87111	Abu87111	Novel	hum
122	18	69.2	1588	5	ABH09437	Abh09437	H. inflie		195	18	69.2	4440	6	ABU87111	Abu87111	Novel	hum
123	18	69.2	1617	8	ADP30660	Adp30660	Human	sec	196	18	69.2	4440	6	ABU91352	Abu91352	Novel	hum
124	18	69.2	1771	8	ADP30674	Adp30674	Human	sec	197	18	69.2	4440	6	ABU90445	Abu90445	Novel	hum
125	18	69.2	1776	8	ADP31298	Adp31298	Human	sec	198	18	69.2	4440	6	ABU97036	Abu97036	Novel	hum
126	18	69.2	1776	9	ADY62715	Ady62715	Human	alp	199	18	69.2	4440	6	ABU05282	Abu05282	Novel	hum
127	18	69.2	1782	8	ADP31391	Adp31391	Human	sec	200	18	69.2	4752	8	ADP30581	Adp30581	Novel	hum
128	18	69.2	1782	8	ADP31270	Adp31270	Human	sec	201	18	69.2	4752	8	ADP30651	Adp30651	Novel	hum
129	18	69.2	1794	8	ADP31176	Adp31176	Human	sec	202	18	69.2	5304	8	ADP30706	Adp30706	Novel	hum
130	18	69.2	1878	8	ADP30740	Adp30740	Human	sec	203	18	69.2	6465	8	ADP31600	Adp31600	Novel	hum
131	18	69.2	1894	6	ABU88253	Abu88253	Novel	hum	204	18	69.2	6729	8	ADP31600	Adp31600	Novel	hum
132	18	69.2	1894	6	ABU90132	Abu90132	Novel	hum	205	18	69.2	7285	6	ABU38280	Abu38280	Novel	hum
133	18	69.2	1894	6	ABU96434	Abu96434	Novel	hum	206	18	69.2	7339	6	AAOI6358	AAOI6358	Novel	hum
134	18	69.2	1894	6	ABU99043	Abu99043	Novel	hum	207	18	69.2	10944	8	ADP31310	Adp31310	Novel	hum
135	18	69.2	1894	6	ABU98258	Abu98258	Novel	hum	208	18	69.2	11328	4	AAU00682	AAU00682	Novel	hum
136	18	69.2	1894	6	ABU91964	Abu91964	Novel	hum	209	17	65.4	14	4	AAU00402	AAU00402	Novel	hum
137	18	69.2	1894	6	ABU85268	Abu85268	Novel	hum	210	17	65.4	22	4	ABBA1096	ABBA1096	Novel	hum
138	18	69.2	1894	6	ABU00407	Abu00407	Novel	hum	211	17	65.4	22	4	AAU34872	AAU34872	Novel	hum
139	18	69.2	1894	6	ABU88958	Abu88958	Novel	hum	212	17	65.4	22	4	AAU4519	AAU4519	Novel	hum
140	18	69.2	1894	6	ABU06454	Abu06454	Novel	hum	213	17	65.4	22	4	ABBA2516	ABBA2516	Novel	hum
141	18	69.2	1894	6	ABU95514	Abu95514	Novel	hum	214	17	65.4	22	4	AAU74756	AAU74756	Novel	hum
142	18	69.2	1894	6	ABU95204	Abu95204	Novel	hum	215	17	65.4	22	4	AAU61952	AAU61952	Novel	hum
143	18	69.2	1894	6	ABU90752	Abu90752	Novel	hum	216	17	65.4	22	4	ABG56539	ABG56539	Novel	hum
144	18	69.2	1894	6	ABU93914	Abu93914	Novel	hum	217	17	65.4	22	5	ABG44553	ABG44553	Novel	hum
145	18	69.2	1894	6	ABU86188	Abu86188	Novel	hum	218	17	65.4	22	4	AAU10165	AAU10165	Novel	hum
146	18	69.2	1894	6	ABU82043	Abu82043	Novel	hum	219	17	65.4	52	4	AAU14688	AAU14688	Novel	hum
147	18	69.2	1894	6	ABU07904	Abu07904	Novel	hum	220	17	65.4	52	4	ABBA3648	ABBA3648	Novel	hum
148	18	69.2	1894	6	ABU94224	Abu94224	Novel	hum	221	17	65.4	52	4	AAU27105	AAU27105	Novel	hum
149	18	69.2	1894	6	ABU00097	Abu00097	Novel	hum	222	17	65.4	52	4	ABBA2846	ABBA2846	Novel	hum
150	18	69.2	1894	6	ABU87108	Abu87108	Novel	hum	223	17	65.4	52	4	ABBI9103	ABBI9103	Novel	hum
151	18	69.2	1894	6	ABU91349	Abu91349	Novel	hum	224	17	65.4	52	4	AAU66822	AAU66822	Novel	hum
152	18	69.2	1894	6	ABU90442	Abu90442	Novel	hum	225	17	65.4	52	4	AAU54421	AAU54421	Novel	hum
153	18	69.2	1894	6	ABU97033	Abu97033	Novel	hum	226	17	65.4	52	4	ABG48490	ABG48490	Novel	hum
154	18	69.2	1894	6	ABO05229	AbO05229	Novel	hum	227	17	65.4	52	4	AAU02411	AAU02411	Novel	hum
155	18	69.2	1917	6	ADU15719	AdU15719	C. elegan		228	17	65.4	52	5	ABG36477	ABG36477	Novel	hum
156	18	69.2	1996	8	ADP31599	Adp31599	Human	sec	229	17	65.4	62	4	AAU15925	AAU15925	Novel	hum
157	18	69.2	2088	8	ADP31178	Adp31178	Human	sec	230	17	65.4	62	4	ABBA4912	ABBA4912	Novel	hum
158	18	69.2	2187	8	ADP30882	Adp30882	Human	sec	231	17	65.4	62	4	AAU28428	AAU28428	Novel	hum
159	18	69.2	2272	8	ADP30669	Adp30669	Human	sec	232	17	65.4	62	4	ABBA29741	ABBA29741	Novel	hum
160	18	69.2	2304	8	ADP31252	Adp31252	Human	sec	233	17	65.4	62	4	ABBA20326	ABBA20326	Novel	hum
161	18	69.2	2307	8	ADP31394	Adp31394	Human	sec	234	17	65.4	62	4	AAU68101	AAU68101	Novel	hum
162	18	69.2	2415	8	ADP31023	Adp31023	Human	sec	235	17	65.4	62	4	AAU55723	AAU55723	Novel	hum
163	18	69.2	2469	8	ADP31031	Adp31031	Human	sec	236	17	65.4	62	4	ABG49751	ABG49751	Novel	hum
164	18	69.2	2542	8	ADP31594	Adp31594	Human	sec	237	17	65.4	62	4	AAU03661	AAU03661	Novel	hum
165	18	69.2	2616	9	ADP31253	Adp31253	Human	sec	238	17	65.4	62	5	ABG37626	ABG37626	Novel	hum
166	18	69.2	2616	9	ABBA4675	AbB44675	N. mening		239	17	65.4	68	4	AAU14223	AAU14223	Novel	hum
167	18	69.2	2980	9	AEA36049	Adp36049	Walze Sta		240	17	65.4	68	4	ABBA3170	ABBA3170	Novel	hum
168	18	69.2	3036	8	ADP31595	Adp31595	Human	sec	241	17	65.4	68	4	AAU26633	AAU26633	Novel	hum
169	18	69.2	3070	8	ADP30710	Adp30710	Human	sec	242	17	65.4	68	4	ABBA27998	ABBA27998	Novel	hum
170	18	69.2	3070	8	ADP30700	Adp30700	Human	sec	243	17	65.4	68	4	ABBA18635	ABBA18635	Novel	hum



244	17	65.4	68	4	AAM66354	Aam66354	Human	bon	317	17	65.4	312	8	ADP31505	Adp31505	Human	sec
245	17	65.4	68	4	AAM53966	Aam53966	Human	bira	318	17	65.4	312	8	ADP31454	Adp31454	Human	sec
246	17	65.4	68	4	ABG48020	Abg48020	Human	liv	319	17	65.4	314	7	ABO82622	AbO82622	Pseudomon	
247	17	65.4	68	4	AAM01955	Aam01955	Peptide	#	320	17	65.4	315	8	ADP31685	Adp31685	Human	sec
248	17	65.4	68	5	ABG36002	Abg36002	Human	pep	321	17	65.4	318	8	ADP31066	Adp31066	Human	sec
249	17	65.4	83	3	AAH11510	Aah11510	SEN	virus	322	17	65.4	333	8	ADP31442	Adp31442	Human	sec
250	17	65.4	96	4	AAU65514	Aau65514	Propionib		323	17	65.4	333	8	ADP31239	Adp31239	Human	sec
251	17	65.4	96	4	AAU66170	Aau66170	Propionib		324	17	65.4	336	8	ADP30470	Adp30470	Human	sec
252	17	65.4	96	6	ABME2033	Abme2033	Propionib		325	17	65.4	337	8	ADP31387	Adp31387	Human	sec
253	17	65.4	96	6	ABME2689	Abme2689	Propionib		326	17	65.4	339	8	ADP30519	Adp30519	Human	sec
254	17	65.4	104	4	ABBE68094	Abbe68094	Drosophil		327	17	65.4	345	8	ADP31065	Adp31065	Human	sec
255	17	65.4	127	8	ADP31622	Adp31622	Human	sec	328	17	65.4	348	8	ADP31374	Adp31374	Human	sec
256	17	65.4	139	9	ADW25813	Adw25813	Scorilas	s	329	17	65.4	349	5	ABE80605	AbE80605	Hirtudin	p
257	17	65.4	139	9	AEA06475	Aea06475	S. spongi		330	17	65.4	354	8	ADP30783	Adp30783	Human	sec
258	17	65.4	139	9	AE872826	Ae872826	Carbohydr		331	17	65.4	354	8	ADP30779	Adp30779	Human	sec
259	17	65.4	142	8	ADP31632	Adp31632	Human	sec	332	17	65.4	354	8	ADP30784	Adp30784	Human	sec
260	17	65.4	149	5	ADK34895	Adk34895	Novel	hum	333	17	65.4	354	8	ADP30778	Adp30778	Human	sec
261	17	65.4	150	7	ABO82982	AbO82982	Pseudomon		334	17	65.4	357	8	ADP31223	Adp31223	Human	sec
262	17	65.4	153	8	ADP30467	Adp30467	Human	sec	335	17	65.4	357	8	ADP30581	Adp30581	Human	sec
263	17	65.4	153	8	ADP31511	Adp31511	Human	sec	336	17	65.4	360	8	ADP31209	Adp31209	Human	sec
264	17	65.4	156	8	ABO71866	AbO71866	Pseudomon		337	17	65.4	366	8	ADP31670	Adp31670	Human	sec
265	17	65.4	156	8	ADP30759	Adp30759	Human	sec	338	17	65.4	369	8	ADP30582	Adp30582	Human	sec
266	17	65.4	165	8	ADP31041	Adp31041	Human	sec	339	17	65.4	378	8	ADP31669	Adp31669	Human	sec
267	17	65.4	168	8	ADP31621	Adp31621	Human	sec	340	17	65.4	381	8	ADP31375	Adp31375	Human	sec
268	17	65.4	168	8	ADP31099	Adp31099	Human	sec	341	17	65.4	381	8	ADP30655	Adp30655	Human	sec
269	17	65.4	168	8	ADP30797	Adp30797	Human	sec	342	17	65.4	383	8	ADP31094	Adp31094	Human	sec
270	17	65.4	168	8	ADP30769	Adp30769	Human	sec	343	17	65.4	384	8	ADP30656	Adp30656	Human	sec
271	17	65.4	168	8	ADP30771	Adp30771	Human	sec	344	17	65.4	390	4	AAU30181	Aau30181	Novel	hum
272	17	65.4	169	7	ABO75555	AbO75555	Pseudomon		345	17	65.4	392	8	ADP31061	Adp31061	Human	sec
273	17	65.4	169	7	ABO71598	AbO71598	Pseudomon		346	17	65.4	392	8	ADP31055	Adp31055	Human	sec
274	17	65.4	172	4	AAU64154	Aau64154	Propionib		347	17	65.4	393	8	ADP31345	Adp31345	Human	sec
275	17	65.4	172	6	ABME6073	Abme6073	Propionib		348	17	65.4	399	8	ADP31313	Adp31313	Human	sec
276	17	65.4	172	8	ADP31682	Adp31682	Human	sec	349	17	65.4	399	8	ADP30578	Adp30578	Human	sec
277	17	65.4	177	8	ADP31496	Adp31496	Human	sec	350	17	65.4	402	8	ADP30742	Adp30742	Human	sec
278	17	65.4	180	8	ADP31496	Adp31496	Human	sec	351	17	65.4	402	8	ADP31262	Adp31262	Human	sec
279	17	65.4	187	3	AAV72094	Aav72094	Human	ser	352	17	65.4	411	8	ADP30729	Adp30729	Human	sec
280	17	65.4	187	5	AAE17922	Aae17922	Human	gen	353	17	65.4	411	7	ABO76350	AbO76350	Pseudomon	
281	17	65.4	187	5	AAE17922	Aae17922	Human	gen	354	17	65.4	423	8	ADP31323	Adp31323	Human	sec
282	17	65.4	189	8	ADP31641	Adp31641	Human	sec	355	17	65.4	423	8	ADP30570	Adp30570	Human	sec
283	17	65.4	191	7	ABO70585	AbO70585	Pseudomon		356	17	65.4	432	8	ADP31365	Adp31365	Human	sec
284	17	65.4	192	8	ADP30573	Adp30573	Human	sec	357	17	65.4	442	8	ADP31134	Adp31134	Human	sec
285	17	65.4	195	8	ADP30590	Adp30590	Human	sec	358	17	65.4	446	6	ABR40850	AbR40850	Ricinus	c
286	17	65.4	200	7	ABO73751	AbO73751	Pseudomon		359	17	65.4	456	8	ADP31224	Adp31224	Human	sec
287	17	65.4	202	7	ABME8006	Abme8006	Rice	abio	360	17	65.4	471	8	ADP31567	Adp31567	Human	sec
288	17	65.4	203	8	ADT76102	Adt76102	Plant	ful	361	17	65.4	471	8	ADP30870	Adp30870	Human	sec
289	17	65.4	204	7	ABO82303	AbO82303	Pseudomon		362	17	65.4	472	8	ADP31222	Adp31222	Human	sec
290	17	65.4	213	8	ADP31348	Adp31348	Human	sec	363	17	65.4	474	8	ADP31157	Adp31157	Human	sec
291	17	65.4	222	7	ABO77741	AbO77741	Pseudomon		364	17	65.4	478	8	ADP31007	Adp31007	Human	sec
292	17	65.4	222	8	ADP31354	Adp31354	Human	sec	365	17	65.4	501	8	ADP31689	Adp31689	Human	sec
293	17	65.4	225	8	ADP30589	Adp30589	Human	sec	366	17	65.4	522	8	ADP31190	Adp31190	Human	sec
294	17	65.4	225	8	ADP31531	Adp31531	Human	sec	367	17	65.4	522	8	ADP31070	Adp31070	Human	sec
295	17	65.4	229	7	ABO80501	AbO80501	Pseudomon		368	17	65.4	528	8	ADP31525	Adp31525	Human	sec
296	17	65.4	243	7	ABO77068	AbO77068	Pseudomon		369	17	65.4	531	8	ADP30594	Adp30594	Human	sec
297	17	65.4	248	8	ADP30662	Adp30662	Human	sec	370	17	65.4	531	8	ADP31696	Adp31696	Human	sec
298	17	65.4	249	7	ADT76961	Adt76961	Novel	hum	371	17	65.4	534	8	ADP30849	Adp30849	Human	sec
299	17	65.4	254	8	ADP30739	Adp30739	Human	sec	372	17	65.4	543	8	ADP31506	Adp31506	Human	sec
300	17	65.4	264	8	ADP30788	Adp30788	Human	sec	373	17	65.4	543	8	ADP31393	Adp31393	Human	sec
301	17	65.4	272	8	ADP30693	Adp30693	Human	sec	374	17	65.4	548	2	AAW40306	Aaw40306	Parapox	v
302	17	65.4	273	8	ADP31236	Adp31236	Human	sec	375	17	65.4	552	8	ADP31524	Adp31524	Human	sec
303	17	65.4	279	8	ADP31509	Adp31509	Human	sec	376	17	65.4	559	8	ADP30716	Adp30716	Human	sec
304	17	65.4	279	8	ADP31503	Adp31503	Human	sec	377	17	65.4	563	8	ADP30881	Adp30881	Human	sec
305	17	65.4	279	8	ADP31671	Adp31671	Human	sec	378	17	65.4	564	8	ADP30743	Adp30743	Human	sec
306	17	65.4	281	7	ABO80453	AbO80453	Pseudomon		379	17	65.4	564	8	ADP31194	Adp31194	Human	sec
307	17	65.4	291	8	ADP31243	Adp31243	Human	sec	380	17	65.4	573	8	ADP31078	Adp31078	Human	sec
308	17	65.4	291	8	ADP31242	Adp31242	Human	sec	381	17	65.4	574	8	ADP31133	Adp31133	Human	sec
309	17	65.4	294	8	ADP31076	Adp31076	Human	sec	382	17	65.4	585	8	ADP31392	Adp31392	Human	sec
310	17	65.4	294	8	ADP31045	Adp31045	Human	sec	383	17	65.4	585	8	ADP31458	Adp31458	Human	sec
311	17	65.4	297	8	ADP31312	Adp31312	Human	sec	384	17	65.4	588	8	ADP30877	Adp30877	Human	sec
312	17	65.4	300	8	ADP31301	Adp31301	Human	sec	385	17	65.4	588	8	ADP31680	Adp31680	Human	sec
313	17	65.4	302	8	ADP31030	Adp31030	Human	sec	386	17	65.4	603	8	ADP31150	Adp31150	Human	sec
314	17	65.4	302	8	ADP31032	Adp31032	Human	sec	387	17	65.4	612	8	ADP31467	Adp31467	Human	sec
315	17	65.4	302	8	ADP31025	Adp31025	Human	sec	388	17	65.4	618	8	ADP31467	Adp31467	Human	sec
316	17	65.4	306	8	ADP31205	Adp31205	Human	sec	389	17	65.4	618	8	ADP31466	Adp31466	Human	sec

390	17	65.4	618	8	ADP31554	Adp31554 Human sec	463	17	65.4	996	8	ADP31538	Adp31538 Human sec
391	17	65.4	621	8	ADP31147	Adp31147 Human sec	464	17	65.4	1017	4	AB59813	Adb59813 TvlD prot
392	17	65.4	624	8	ADP31324	Adp31324 Human sec	465	17	65.4	1017	8	ADP31268	Adp31268 Human sec
393	17	65.4	624	8	ADP31325	Adp31325 Human sec	466	17	65.4	1035	8	ADP31552	Adp31552 Human sec
394	17	65.4	629	8	ADP31295	Adp31295 Human sec	467	17	65.4	1041	8	ADP30768	Adp30768 Human sec
395	17	65.4	639	8	ADP30861	Adp30861 Human sec	468	17	65.4	1044	8	ADP31517	Adp31517 Human sec
396	17	65.4	639	9	AD210464	Ad210464 P. gingiv	469	17	65.4	1050	8	ADP31370	Adp31370 Human sec
397	17	65.4	642	8	ADP31161	Adp31161 Human sec	470	17	65.4	1053	8	ADP30886	Adp30886 Human sec
398	17	65.4	645	8	ADP31124	Adp31124 Human sec	471	17	65.4	1056	8	ADP31082	Adp31082 Human sec
399	17	65.4	645	8	ADP31125	Adp31125 Human sec	472	17	65.4	1059	8	ADP31042	Adp31042 Human sec
400	17	65.4	654	8	ADP31002	Adp31002 Human sec	473	17	65.4	1062	8	ADP31369	Adp31369 Human sec
401	17	65.4	660	8	ADP31606	Adp31606 Human sec	474	17	65.4	1066	8	ADP30563	Adp30563 Human sec
402	17	65.4	666	8	ADP30466	Adp30466 Human sec	475	17	65.4	1076	6	AB081145	Abu81145 Human PRO
403	17	65.4	669	8	ADP31598	Adp31598 Human sec	476	17	65.4	1076	6	ABU66845	Abu66845 Human PRO
404	17	65.4	669	8	ADP31493	Adp31493 Human sec	477	17	65.4	1083	8	ADP30473	Adp30473 Human sec
405	17	65.4	670	8	ADP31001	Adp31001 Human sec	478	17	65.4	1086	8	ADP31175	Adp31175 Human sec
406	17	65.4	672	8	ADP30498	Adp30498 Human sec	479	17	65.4	1086	8	ADP31447	Adp31447 Human sec
407	17	65.4	681	8	ADP31053	Adp31053 Human sec	480	17	65.4	1086	8	ADP31629	Adp31629 Human sec
408	17	65.4	687	8	ADP31364	Adp31364 Human sec	481	17	65.4	1087	8	ADP30978	Adp30978 Human sec
409	17	65.4	705	8	ADP31623	Adp31623 Human sec	482	17	65.4	1104	8	ADP31156	Adp31156 Human sec
410	17	65.4	711	8	ADP31215	Adp31215 Human sec	483	17	65.4	1113	8	ADP31508	Adp31508 Human sec
411	17	65.4	711	8	ADP31535	Adp31535 Human sec	484	17	65.4	1116	8	ADP31203	Adp31203 Human sec
412	17	65.4	714	8	ADP31561	Adp31561 Human sec	485	17	65.4	1116	8	ADP31202	Adp31202 Human sec
413	17	65.4	717	8	ADP31410	Adp31410 Human sec	486	17	65.4	1128	8	ADP30483	Adp30483 Human sec
414	17	65.4	725	8	ADP31092	Adp31092 Human sec	487	17	65.4	1134	8	ADP31537	Adp31537 Human sec
415	17	65.4	732	8	ADP30650	Adp30650 Human sec	488	17	65.4	1141	8	ADP30677	Adp30677 Human sec
416	17	65.4	746	8	ADP30982	Adp30982 Human sec	489	17	65.4	1147	8	ADP30965	Adp30965 Human sec
417	17	65.4	750	8	ADP31131	Adp31131 Human sec	490	17	65.4	1155	8	ADP31455	Adp31455 Human sec
418	17	65.4	755	8	ADP31559	Adp31559 Human sec	491	17	65.4	1168	8	ADP31046	Adp31046 Human sec
419	17	65.4	756	8	ADP31039	Adp31039 Human sec	492	17	65.4	1189	8	ADP31043	Adp31043 Human sec
420	17	65.4	764	8	ADP31614	Adp31614 Human sec	493	17	65.4	1191	8	ADP30993	Adp30993 Human sec
421	17	65.4	765	8	ADP31149	Adp31149 Human sec	494	17	65.4	1192	8	ADP31180	Adp31180 Human sec
422	17	65.4	774	8	ADP30885	Adp30885 Human sec	495	17	65.4	1199	8	ADP31044	Adp31044 Human sec
423	17	65.4	774	8	ADP31373	Adp31373 Human sec	496	17	65.4	1200	8	ADP31603	Adp31603 Human sec
424	17	65.4	779	8	ADP30915	Adp30915 Human sec	497	17	65.4	1221	8	ADP31293	Adp31293 Human sec
425	17	65.4	779	8	ADP30897	Adp30897 Human sec	498	17	65.4	1224	8	ADP31426	Adp31426 Human sec
426	17	65.4	780	8	ADP31077	Adp31077 Human sec	499	17	65.4	1227	8	ADP31602	Adp31602 Human sec
427	17	65.4	783	8	ADP31398	Adp31398 Human sec	500	17	65.4	1227	8	ADP31210	Adp31210 Human sec
428	17	65.4	804	8	ADP31635	Adp31635 Human sec	501	17	65.4	1248	8	ADP31346	Adp31346 Human sec
429	17	65.4	806	8	ADP31110	Adp31110 Human sec	502	17	65.4	1252	8	ADP30678	Adp30678 Human sec
430	17	65.4	813	8	ADP30561	Adp30561 Human sec	503	17	65.4	1260	8	ADP31533	Adp31533 Human sec
431	17	65.4	816	8	ADP31249	Adp31249 Human sec	504	17	65.4	1269	8	ADP31382	Adp31382 Human sec
432	17	65.4	816	8	ADP31581	Adp31581 Human sec	505	17	65.4	1269	8	ADP31381	Adp31381 Human sec
433	17	65.4	823	4	ABG25826	Abg25826 Novel hum	506	17	65.4	1282	8	ADP31328	Adp31328 Human sec
434	17	65.4	827	8	ADN22420	Adn22420 Bacterial	507	17	65.4	1288	8	ADP31114	Adp31114 Human sec
435	17	65.4	831	8	ADP31333	Adp31333 Human sec	508	17	65.4	1289	8	ADP30675	Adp30675 Human sec
436	17	65.4	831	8	ADP31179	Adp31179 Human sec	509	17	65.4	1300	6	ABU88254	Abu88254 Novel hum
437	17	65.4	876	8	ADP31220	Adp31220 Human sec	510	17	65.4	1300	6	ABU90133	Abu90133 Novel hum
438	17	65.4	887	8	ADP30554	Adp30554 Human sec	511	17	65.4	1300	6	ABU96435	Abu96435 Novel hum
439	17	65.4	887	8	ADP30548	Adp30548 Human sec	512	17	65.4	1300	6	ABU99044	Abu99044 Novel hum
440	17	65.4	889	8	ADP31648	Adp31648 Human sec	513	17	65.4	1300	6	ABU98259	Abu98259 Novel hum
441	17	65.4	890	8	ADP31059	Adp31059 Human sec	514	17	65.4	1300	6	ABU91965	Abu91965 Novel hum
442	17	65.4	897	8	ADP30914	Adp30914 Human sec	515	17	65.4	1300	6	ABU85269	Abu85269 Novel hum
443	17	65.4	903	8	ADP31294	Adp31294 Human sec	516	17	65.4	1300	6	ABO00408	Abu00408 Novel hum
444	17	65.4	906	6	ADP31573	Adp31573 Human sec	517	17	65.4	1300	6	ABU88959	Abu88959 Novel hum
445	17	65.4	908	6	ABU08492	Abu08492 ALpha-hel	518	17	65.4	1300	6	ABO06455	Abu06455 Novel hum
446	17	65.4	912	8	ADP31450	Adp31450 Human sec	519	17	65.4	1300	6	ABU95515	Abu95515 Novel hum
447	17	65.4	918	8	ADP31569	Adp31569 Human sec	520	17	65.4	1300	6	ABU95205	Abu95205 Novel hum
448	17	65.4	921	8	ADP31522	Adp31522 Human sec	521	17	65.4	1300	6	ABU90753	Abu90753 Novel hum
449	17	65.4	925	5	AAO14246	Aao14246 Human sec	522	17	65.4	1300	6	ABU93915	Abu93915 Novel hum
450	17	65.4	933	8	ADP31510	Adp31510 Human sec	523	17	65.4	1300	6	ABU86189	Abu86189 Novel hum
451	17	65.4	936	8	ADP31597	Adp31597 Human sec	524	17	65.4	1300	6	ABU82044	Abu82044 Novel hum
452	17	65.4	939	8	ADP31086	Adp31086 Human sec	525	17	65.4	1300	6	ABU07905	Abu07905 Novel hum
453	17	65.4	939	8	ADP30726	Adp30726 Human sec	526	17	65.4	1300	6	ABU94225	Abu94225 Novel hum
454	17	65.4	942	8	ADN22417	Adn22417 Bacterial	527	17	65.4	1300	6	ABO00098	Abu00098 Novel hum
455	17	65.4	948	8	ADP30586	Adp30586 Human sec	528	17	65.4	1300	6	ABU87109	Abu87109 Novel hum
456	17	65.4	951	8	ADP31413	Adp31413 Human sec	529	17	65.4	1300	6	ABU91350	Abu91350 Novel hum
457	17	65.4	966	8	ADP30745	Adp30745 Human sec	530	17	65.4	1300	6	ABU90443	Abu90443 Novel hum
458	17	65.4	967	8	ADP30698	Adp30698 Human sec	531	17	65.4	1300	6	ABU97034	Abu97034 Novel hum
459	17	65.4	967	8	ADP30718	Adp30718 Human sec	532	17	65.4	1302	8	ADP31695	Adp31695 Human sec
460	17	65.4	976	8	ADP30688	Adp30688 Human sec	533	17	65.4	1302	8	ADP30751	Adp30751 Human sec
461	17	65.4	981	8	ADP30547	Adp30547 Human sec	534	17	65.4	1303	8	ADP30751	Adp30751 Human sec
462	17	65.4	990	8	ADP31553	Adp31553 Human sec	535	17	65.4	1312	8	ADP30999	Adp30999 Human sec

536	17	65.4	1314	8	ADP31197	Adp31197	Human	sec	609	17	65.4	1803	8	ADP30562	Adp30562	Human	sec
537	17	65.4	1344	8	ADP31211	Adp31211	Human	sec	610	17	65.4	1815	8	ADP31601	Adp31601	Human	sec
538	17	65.4	1365	8	ADP31661	Adp31661	Human	sec	611	17	65.4	1827	8	ADP31170	Adp31170	Human	sec
539	17	65.4	1371	8	ADP31646	Adp31646	Human	sec	612	17	65.4	1833	8	ADP30642	Adp30642	Human	sec
540	17	65.4	1380	8	ADP31566	Adp31566	Human	sec	613	17	65.4	1848	8	ADP31372	Adp31372	Human	sec
541	17	65.4	1386	8	ADP31371	Adp31371	Human	sec	614	17	65.4	1849	8	ADP30752	Adp30752	Human	sec
542	17	65.4	1398	8	ADP31488	Adp31488	Human	sec	615	17	65.4	1875	8	ADP31664	Adp31664	Human	sec
543	17	65.4	1401	8	ADP30532	Adp30532	Human	sec	616	17	65.4	1933	8	ADP30889	Adp30889	Human	sec
544	17	65.4	1417	8	ADP31160	Adp31160	Human	sec	617	17	65.4	1933	8	ADP30902	Adp30902	Human	sec
545	17	65.4	1437	8	ADP31357	Adp31357	Human	sec	618	17	65.4	1933	8	ADP31662	Adp31662	Human	sec
546	17	65.4	1454	8	ADP31177	Adp31177	Human	sec	619	17	65.4	1968	8	ADP30689	Adp30689	Human	sec
547	17	65.4	1456	8	ADP30923	Adp30923	Human	sec	620	17	65.4	2001	8	ADP31644	Adp31644	Human	sec
548	17	65.4	1464	8	ADP31040	Adp31040	Human	sec	621	17	65.4	2020	8	ADP31056	Adp31056	Human	sec
549	17	65.4	1464	8	ADP31437	Adp31437	Human	sec	622	17	65.4	2027	8	ADP31058	Adp31058	Human	sec
550	17	65.4	1485	8	ADP31383	Adp31383	Human	sec	623	17	65.4	2052	9	AEH87653	AEh87653	Human	ino
551	17	65.4	1485	8	ADP31384	Adp31384	Human	sec	624	17	65.4	2058	8	ADP31630	Adp31630	Human	sec
552	17	65.4	1488	8	ADP31385	Adp31385	Human	sec	625	17	65.4	2123	8	ADP30657	Adp30657	Human	sec
553	17	65.4	1488	8	ADP31386	Adp31386	Human	sec	626	17	65.4	2123	8	ADP30564	Adp30564	Human	sec
554	17	65.4	1494	8	ADP31650	Adp31650	Human	sec	627	17	65.4	2127	8	ADP31327	Adp31327	Human	sec
555	17	65.4	1506	8	ADP30596	Adp30596	Human	sec	628	17	65.4	2148	8	ADP30974	Adp30974	Human	sec
556	17	65.4	1521	8	ADP30542	Adp30542	Human	sec	629	17	65.4	2260	8	ADP30687	Adp30687	Human	sec
557	17	65.4	1521	8	ADP30537	Adp30537	Human	sec	630	17	65.4	2272	8	ADP31136	Adp31136	Human	sec
558	17	65.4	1530	8	ADP31536	Adp31536	Human	sec	631	17	65.4	2349	8	ADP30959	Adp30959	Human	sec
559	17	65.4	1536	8	ADP31302	Adp31302	Human	sec	632	17	65.4	2382	8	ADP31341	Adp31341	Human	sec
560	17	65.4	1539	8	ADP31201	Adp31201	Human	sec	633	17	65.4	2454	8	ADP30469	Adp30469	Human	sec
561	17	65.4	1560	8	ADP31200	Adp31200	Human	sec	634	17	65.4	2484	8	ADP66690	ADp66690	Human	mis
562	17	65.4	1560	8	ADP31570	Adp31570	Human	sec	635	17	65.4	2508	6	ADA15721	ADA15721	C. elegans	
563	17	65.4	1560	8	ADP31139	Adp31139	Human	sec	636	17	65.4	2535	8	ADP31146	Adp31146	Human	sec
564	17	65.4	1560	8	ADP31631	Adp31631	Human	sec	637	17	65.4	2544	6	ADA15717	ADA15717	C. elegans	
565	17	65.4	1566	8	ADP31407	Adp31407	Human	sec	638	17	65.4	2547	8	ADP31665	Adp31665	Human	sec
566	17	65.4	1584	8	ADP31405	Adp31405	Human	sec	639	17	65.4	2601	6	ADA15723	ADA15723	C. elegans	
567	17	65.4	1587	8	ADP30591	Adp30591	Human	sec	640	17	65.4	2622	8	ADP31663	Adp31663	Human	sec
568	17	65.4	1596	8	ADP31491	Adp31491	Human	sec	641	17	65.4	2664	8	ADP31546	Adp31546	Human	sec
569	17	65.4	1635	4	AAH59826	AAh59826	Protein	#	642	17	65.4	2688	8	ADN11593	ADn11593	Human	CD9
570	17	65.4	1633	4	ADP30552	Adp30552	Human	sec	643	17	65.4	2724	8	ADP31233	Adp31233	Human	sec
571	17	65.4	1631	4	ABG22481	ABg22481	Novel	hum	644	17	65.4	2790	8	ADP31232	Adp31232	Human	sec
572	17	65.4	1647	8	ADP30670	Adp30670	Human	sec	645	17	65.4	2808	8	ADP31258	Adp31258	Human	sec
573	17	65.4	1647	8	ADP31052	Adp31052	Human	sec	646	17	65.4	2827	8	ADP30681	Adp30681	Human	sec
574	17	65.4	1652	6	ADA15715	AdA15715	C. elegans		647	17	65.4	2828	8	ADP30938	Adp30938	Human	sec
575	17	65.4	1654	8	ADP30997	Adp30997	Human	sec	648	17	65.4	2833	8	ADP31299	Adp31299	Human	sec
576	17	65.4	1662	8	ADP31419	Adp31419	Human	sec	649	17	65.4	2835	8	ADP30572	Adp30572	Human	sec
577	17	65.4	1662	8	ADP31513	Adp31513	Human	sec	650	17	65.4	3046	8	ADQ10419	ADq10419	Human	pol
578	17	65.4	1665	8	ADP31187	Adp31187	Human	sec	651	17	65.4	3144	8	ADP31544	Adp31544	Human	sec
579	17	65.4	1700	3	AAH18144	AAh18144	Plasmodi		652	17	65.4	3201	8	ADP31545	Adp31545	Human	sec
580	17	65.4	1725	8	ADP30654	Adp30654	Human	sec	653	17	65.4	3398	9	AEH87634	AEh87634	Human	ino
581	17	65.4	1737	8	ADP31292	Adp31292	Human	sec	654	17	65.4	3407	8	ADP31060	Adp31060	Human	sec
582	17	65.4	1743	6	ABH88255	ABh88255	Novel	hum	655	17	65.4	3407	8	ADP31062	Adp31062	Human	sec
583	17	65.4	1743	6	ABU90134	ABu90134	Novel	hum	656	17	65.4	3465	8	ADP31234	Adp31234	Human	sec
584	17	65.4	1743	6	ABU96436	ABu96436	Novel	hum	657	17	65.4	3579	8	ADP31098	Adp31098	Human	sec
585	17	65.4	1743	6	ABU99045	ABu99045	Novel	hum	658	17	65.4	3638	8	ADP30981	Adp30981	Human	sec
586	17	65.4	1743	6	ABU98260	ABu98260	Novel	hum	659	17	65.4	4683	8	ADP31260	Adp31260	Human	sec
587	17	65.4	1743	6	ABU91966	ABu91966	Novel	hum	660	17	65.4	4848	8	ADP31259	Adp31259	Human	sec
588	17	65.4	1743	6	ABU85270	ABu85270	Novel	hum	661	17	65.4	5397	8	ADP31068	Adp31068	Human	sec
589	17	65.4	1743	6	ABO00409	ABo00409	Novel	hum	662	17	65.4	5514	8	ADP31186	Adp31186	Human	sec
590	17	65.4	1743	6	ABU88960	ABu88960	Novel	hum	663	17	65.4	5514	8	ADP31591	Adp31591	Human	sec
591	17	65.4	1743	6	ABO06456	ABo06456	Novel	hum	664	17	65.4	5820	8	ADP31118	Adp31118	Human	sec
592	17	65.4	1743	6	ABU95516	ABu95516	Novel	hum	665	17	65.4	8973	8	ADP31119	Adp31119	Human	sec
593	17	65.4	1743	6	ABU95206	ABu95206	Novel	hum	666	17	65.4	8976	8	ADP31125	Adp31125	Human	sec
594	17	65.4	1743	6	ABU90754	ABu90754	Novel	hum	667	17	65.4	9195	8	ADP31494	Adp31494	Human	sec
595	17	65.4	1743	6	ABU93916	ABu93916	Novel	hum	668	17	65.4	10	2	AAE71548	AAe71548	AIDS	viru
596	17	65.4	1743	6	ABU86190	ABu86190	Novel	hum	669	17	65.4	10	1	AAE28918	AAe28918	Peptide	f
597	17	65.4	1743	6	ABU82045	ABu82045	Novel	hum	670	17	65.4	22	5	ABH05791	ABh05791	HIV	tat
598	17	65.4	1743	6	ABU07906	ABu07906	Novel	hum	671	17	65.4	22	9	ADY76485	ADy76485	Hepatitis	
599	17	65.4	1743	6	ABU94226	ABu94226	Novel	hum	672	17	65.4	25	5	AAE21859	AAe21859	Hepatitis	
600	17	65.4	1743	6	ABO00099	ABo00099	Novel	hum	673	17	65.4	25	5	AAE19915	AAe19915	Hepatitis	
601	17	65.4	1743	6	ABU87110	ABu87110	Novel	hum	674	17	65.4	25	7	AEW00366	AEw00366	Hepatitis	
602	17	65.4	1743	6	ABU91351	ABu91351	Novel	hum	675	17	65.4	35	8	ADG47681	ADg47681	HCV	NS3/4
603	17	65.4	1743	6	ABU90444	ABu90444	Novel	hum	676	17	65.4	34	5	AAH87022	AAh87022	Adenoviru	
604	17	65.4	1743	6	ABU97035	ABu97035	Novel	hum	677	17	65.4	41	4	ABH41607	ABh41607	Peptide	#
605	17	65.4	1743	6	ABO05231	ABo05231	Novel	hum	678	17	65.4	41	4	AAH35400	AAh35400	Protein	#
606	17	65.4	1749	8	ADP31408	Adp31408	Human	sec	679	17	65.4	41	4	ABH25433	ABh25433	Protein	#
607	17	65.4	1755	8	ADP31446	Adp31446	Human	sec	680	17	65.4	41	4	AAW75287	AAw75287	Human	Don
608	17	65.4	1789	8	ADP30962	Adp30962	Human	sec	681	17	61.5	41	4	AAW62478	AAw62478	Human	Bra

682	16	61.5	41	4	ABG57048	Abg57048 Human 11v	755	16	61.5	110	6	ABM65599	Abm65599 Proponiob
683	16	61.5	41	5	ABG44927	Abg44927 Human pep	756	16	61.5	111	5	ADK35517	Adk35517 Novel hum
684	16	61.5	44	2	AAR98208	Aar98208 Nucleicid	757	16	61.5	112	7	ABO68317	AbO68317 Pseudom
685	16	61.5	45	8	ADT92777	Adt92777 Mouse Cal	758	16	61.5	112	8	ADX88804	Adx88804 Plant ful
686	16	61.5	45	8	ADT92788	Adt92788 Mouse Cal	759	16	61.5	113	8	ABT03134	AbT03134 SARS coro
687	16	61.5	47	9	ADZ47821	Adz47821 Amino aci	760	16	61.5	113	9	AEB91569	Aeb91569 Microbial
688	16	61.5	48	3	AA657161	AA657161 Arabidops	761	16	61.5	114	5	ABP06472	Abp06472 Human ORF
689	16	61.5	48	3	AA657161	AA657161 Arabidops	762	16	61.5	115	4	AAH83845	Aah83845 Human lmm
690	16	61.5	48	9	ADV76464	Adv76464 Human HER	763	16	61.5	116	8	ADP30955	Adp30955 Human sec
691	16	61.5	48	9	AEA34233	Aea34233 Optum pop	764	16	61.5	118	4	ABG22738	Abg22738 Novel hum
692	16	61.5	51	8	ADP29386	Adp29386 Human sec	765	16	61.5	121	9	ABM96843	Abm96843 M. xanthu
693	16	61.5	51	8	ADP29383	Adp29383 Human sec	766	16	61.5	122	4	ABBI1761	Abbi1761 Human ner
694	16	61.5	52	5	ABP31338	Abp31338 Human ORF	767	16	61.5	122	4	ABO83310	AbO83310 Pseudom
695	16	61.5	58	4	AAU48534	Aua48534 Proponiob	768	16	61.5	123	4	AAH90307	Aah90307 Human lmm
696	16	61.5	58	6	ABM45053	Abm45053 Proponiob	769	16	61.5	125	4	ABG04861	Abg04861 Novel hum
697	16	61.5	58	6	ABM45053	Abm45053 Proponiob	770	16	61.5	133	4	AAU64082	Aau64082 Proponiob
698	16	61.5	61	4	AAH19250	Aah19250 Peptide #	771	16	61.5	133	6	ABM60601	Abm60601 Human sec
699	16	61.5	61	4	ABB38496	Abb38496 Peptide #	772	16	61.5	133	8	ADP31069	Adp31069 Human sec
700	16	61.5	61	4	AAH31941	Aah31941 Peptide #	773	16	61.5	134	7	ABO75321	AbO75321 Pseudom
701	16	61.5	61	4	ABB23648	Abb23648 Protein #	774	16	61.5	136	7	ABO73061	AbO73061 Pseudom
702	16	61.5	61	4	AAH71647	Aam71647 Human bon	775	16	61.5	137	4	AAH93766	Aah93766 Human pro
703	16	61.5	61	4	AAH59112	Aam59112 Human bra	776	16	61.5	137	7	ABO82722	AbO82722 Pseudom
704	16	61.5	61	5	ABG53331	Abg53331 Human 11v	777	16	61.5	137	7	ABO80767	AbO80767 Pseudom
705	16	61.5	63	4	AAU59447	Aau59447 Human pep	778	16	61.5	139	7	ABO72721	AbO72721 Pseudom
706	16	61.5	63	6	ABM55966	Abm55966 Proponiob	779	16	61.5	139	7	ABO76095	AbO76095 Pseudom
707	16	61.5	63	6	ABR63898	AbR63898 P. furiosus	780	16	61.5	140	7	ABO83055	AbO83055 Pseudom
708	16	61.5	65	4	ABP35884	Abp35884 CRRB_GADM	781	16	61.5	140	7	ABG33884	Abg33884 Human pro
709	16	61.5	65	4	ABP35883	Abp35883 CTRR_GADM	782	16	61.5	141	4	ABO84154	AbO84154 Pseudom
710	16	61.5	65	4	AAU41244	Aau41244 Proponiob	783	16	61.5	146	7	ABO72664	AbO72664 Pseudom
711	16	61.5	65	6	ABM37763	Abm37763 Proponiob	784	16	61.5	146	7	ABG29376	AbG29376 Human gen
712	16	61.5	66	8	AAAG57014	Aag57014 Arabidops	785	16	61.5	147	4	ABG29376	AbG29376 Human gen
713	16	61.5	68	8	ADP31694	Adp31694 Human sec	786	16	61.5	150	2	AAH20620	Aah20620 PTM3 huma
714	16	61.5	70	7	ABO73323	AbO73323 Pseudom	787	16	61.5	150	3	AAH20620	Aah20620 PTM3 huma
715	16	61.5	71	4	AAH14601	Aam14601 Peptide #	788	16	61.5	150	7	ABO64648	AbO64648 Human sec
716	16	61.5	71	4	ABB33561	Abb33561 Peptide #	789	16	61.5	150	7	ADP31605	Adp31605 Human SAR
717	16	61.5	71	4	AAH27022	Aam27022 Peptide #	790	16	61.5	152	8	ADV95514	Adv95514 Human pol
718	16	61.5	71	4	ABB28383	Abb28383 Peptide #	791	16	61.5	153	8	ABP62822	Abp62822 Human CGD
719	16	61.5	71	4	ABBI19018	Abbi19018 Protein #	792	16	61.5	154	5	AAH36181	Aah36181 Human CGD
720	16	61.5	71	4	AAH66735	Aam66735 Human bon	793	16	61.5	155	7	ABO77012	AbO77012 Pseudom
721	16	61.5	71	4	AAU53020	Aau53020 Proponiob	794	16	61.5	155	7	ABO83103	AbO83103 Pseudom
722	16	61.5	71	4	AAH54335	Aam54335 Human bra	795	16	61.5	156	7	ABO83103	AbO83103 Pseudom
723	16	61.5	71	4	ABG48403	Abg48403 Human 11v	796	16	61.5	158	7	ABO76328	AbO76328 Pseudom
724	16	61.5	71	4	AAH02326	Aam02326 Peptide #	797	16	61.5	158	7	ABO79744	AbO79744 Human gen
725	16	61.5	71	5	ABG35389	Abg35389 Human pep	798	16	61.5	158	8	ABO59415	AbO59415 Human gen
726	16	61.5	71	6	ABM49539	Abm49539 Proponiob	799	16	61.5	159	4	AAO10454	Aao10454 Human pol
727	16	61.5	71	9	ADZ47683	Adz47683 Ant1-1L-1	800	16	61.5	159	7	ABO79796	AbO79796 Pseudom
728	16	61.5	74	4	AAU47686	Aau47686 Proponiob	801	16	61.5	160	7	ABO73965	AbO73965 Pseudom
729	16	61.5	74	6	ABM44205	Abm44205 Human bon	802	16	61.5	162	3	AAH95932	Aah95932 Porcine a
730	16	61.5	80	4	AAH70800	Aam70800 Human bon	803	16	61.5	162	7	ABO69461	AbO69461 Pseudom
731	16	61.5	80	4	AAH59526	Aam59526 Proponiob	804	16	61.5	162	7	ABO76207	AbO76207 Pseudom
732	16	61.5	80	6	ABM65045	Abm65045 Proponiob	805	16	61.5	162	7	ABO79469	AbO79469 Human sec
733	16	61.5	88	4	AAH74029	Aam74029 Human col	806	16	61.5	162	8	ADP30813	Adp30813 Human sec
734	16	61.5	90	8	ADP31399	Adp31399 Human sec	807	16	61.5	162	8	ADH88496	Adh88496 Plant ful
735	16	61.5	90	8	ADP31400	Adp31400 Human sec	808	16	61.5	165	4	AAU54621	Aau54621 Proponiob
736	16	61.5	92	4	AAU46180	Aau46180 Proponiob	809	16	61.5	165	6	ABM51140	Abm51140 Proponiob
737	16	61.5	92	6	ABM42699	Abm42699 Proponiob	810	16	61.5	165	8	ADP31286	Adp31286 Human sec
738	16	61.5	93	4	ABBA4737	Abba4737 Peptide #	811	16	61.5	167	2	AAH38569	Aah38569 Neisseria
739	16	61.5	93	4	AAH37635	Aam37635 Peptide #	812	16	61.5	167	9	AEH84987	Aeh84987 N. mening
740	16	61.5	93	4	AAH77475	Aam77475 Human bon	813	16	61.5	168	7	ABO80868	AbO80868 Pseudom
741	16	61.5	93	5	ABG46494	Abg46494 Human pep	814	16	61.5	168	8	ADP30841	Adp30841 Human sec
742	16	61.5	93	5	ADP30859	Adp30859 Human sec	815	16	61.5	171	8	ABO60389	AbO60389 Human gen
743	16	61.5	95	8	ADG23362	Adg22362 Cyanophag	816	16	61.5	171	8	ADP30599	Adp30599 Human sec
744	16	61.5	104	2	AAH12843	Aah12843 HTLV-1 pr	817	16	61.5	173	7	ABO74910	AbO74910 Pseudom
745	16	61.5	104	3	AAH24831	Aah24831 Arabidops	818	16	61.5	174	8	ADP30801	Adp30801 Human sec
746	16	61.5	104	6	AAU55097	Aau55097 Proponiob	819	16	61.5	175	4	ABH68003	Abh68003 Drosophil
747	16	61.5	104	6	ABM51616	Abm51616 Proponiob	820	16	61.5	180	8	ADP30820	Adp30820 Human sec
748	16	61.5	104	7	ABO78018	AbO78018 Pseudom	821	16	61.5	180	8	ADP30825	Adp30825 Human sec
749	16	61.5	105	4	AAU42091	Aau42091 Proponiob	822	16	61.5	180	8	ADP30828	Adp30828 Human sec
750	16	61.5	105	6	ABM38610	Abm38610 Proponiob	823	16	61.5	180	8	ADP30821	Adp30821 Human sec
751	16	61.5	106	6	AAU58682	Aau58682 Proponiob	824	16	61.5	180	8	ADP30826	Adp30826 Human sec
752	16	61.5	106	6	ABM55201	Abm55201 Proponiob	825	16	61.5	180	8	ADP30827	Adp30827 Human sec
753	16	61.5	108	4	AAU42243	Aau42243 Proponiob	826	16	61.5	181	8	ADP31029	Adp31029 Human sec
754	16	61.5	108	6	ABM38762	Abm38762 Proponiob	827	16	61.5	181	8	ADP31029	Adp31029 Human sec

828	16	61.5	181	8	ADQ67035	Adq67035	Novel hum	901	16	61.5	266	9	AEb42329	AEb42329	L. pneumo
829	16	61.5	185	7	ABb11163	Abb11163	Human orp	902	16	61.5	267	8	ADp30822	ADp30822	Human sec
830	16	61.5	186	4	ABO68891	ABO68891	Pseudomon	903	16	61.5	267	8	ADp31483	ADp31483	Human sec
831	16	61.5	186	8	ADp30789	Adp30789	Human sec	904	16	61.5	268	8	ADp31697	ADp31697	Human sec
832	16	61.5	186	8	ADp30787	Adp30787	Human sec	905	16	61.5	270	8	ADp31435	ADp31435	Human sec
833	16	61.5	186	8	ADp30786	Adp30786	Human sec	906	16	61.5	272	5	AAO20952	AAO20952	Human-N-K
834	16	61.5	186	8	ADp30785	Adp30785	Human sec	907	16	61.5	272	6	ABg76399	ABg76399	Antno act
835	16	61.5	187	7	ABO68818	ABO68818	Pseudomon	908	16	61.5	273	8	ADp31336	ADp31336	Human sec
836	16	61.5	188	7	ADp306439	Adp306439	Plant ful	909	16	61.5	273	8	ADp31583	ADp31583	Human sec
837	16	61.5	189	8	ADp30641	Adp30641	Human sec	910	16	61.5	274	8	ADp31679	ADp31679	Human sec
838	16	61.5	192	8	ADp31335	Adp31335	Human sec	911	16	61.5	275	8	ADp65898	ADp65898	Novel hum
839	16	61.5	192	8	ADp31576	Adp31576	Human sec	912	16	61.5	275	8	ADp60000	ADp60000	Plant pol
840	16	61.5	193	7	ABO73977	ABO73977	Pseudomon	913	16	61.5	277	8	ADp30503	ADp30503	Human sec
841	16	61.5	195	8	ADp30696	Adp30696	Human sec	914	16	61.5	278	3	AAO66278	AAO66278	Arabidops
842	16	61.5	197	7	ABO73630	ABO73630	Pseudomon	915	16	61.5	278	7	ABO68255	ABO68255	Pseudomon
843	16	61.5	198	8	ADp30477	Adp30477	Human sec	916	16	61.5	279	3	AAO32271	AAO32271	Arabidops
844	16	61.5	198	8	ADp30481	Adp30481	Human sec	917	16	61.5	279	3	AAO66469	AAO66469	Arabidops
845	16	61.5	198	8	ADp30840	Adp30840	Human sec	918	16	61.5	279	8	ADp31553	ADp31553	Human sec
846	16	61.5	200	3	AAO54845	AAO54845	Arabidops	919	16	61.5	279	8	ADp31107	ADp31107	Human sec
847	16	61.5	201	8	ADp30711	Adp30711	Human sec	920	16	61.5	280	8	ADp66570	ADp66570	Novel hum
848	16	61.5	201	8	ADp31351	Adp31351	Human sec	921	16	61.5	280	8	ADp309856	ADp309856	Human pro
849	16	61.5	203	8	ADp66375	ADp66375	Novel hum	922	16	61.5	282	7	ADp32143	ADp32143	Human hai
850	16	61.5	204	8	ADp31404	Adp31404	Human sec	923	16	61.5	285	3	AAO68088	AAO68088	Arabidops
851	16	61.5	209	4	ABb70477	Abb70477	Drosoph11	924	16	61.5	286	7	ABO66130	ABO66130	Klebslell
852	16	61.5	210	8	ADp58395	ADp58395	Plant pol	925	16	61.5	286	8	ADp31463	ADp31463	Human sec
853	16	61.5	210	8	ADp66008	ADp66008	Plant ful	926	16	61.5	291	8	ADp31193	ADp31193	Human sec
854	16	61.5	212	7	ABO79141	ABO79141	Pseudomon	927	16	61.5	297	8	ADp30744	ADp30744	Human sec
855	16	61.5	213	8	ADp31339	Adp31339	Human sec	928	16	61.5	298	8	ADp43100	ADp43100	Plant tra
856	16	61.5	214	8	ABM84290	ABM84290	Human sec	929	16	61.5	300	8	ADp30616	ADp30616	Human sec
857	16	61.5	215	7	ADpE08255	ADpE08255	Novel pro	930	16	61.5	300	8	ADp30775	ADp30775	Human sec
858	16	61.5	216	4	AAO52527	AAO52527	Human col	931	16	61.5	309	2	AAO37650	AAO37650	Chlamydia
859	16	61.5	217	8	ADpY10185	ADpY10185	Plant ful	932	16	61.5	310	8	ADp30521	ADp30521	Human sec
860	16	61.5	218	7	ABO67270	ABO67270	Klebslell	933	16	61.5	311	3	AAO66277	AAO66277	Arabidops
861	16	61.5	219	8	ADp30593	Adp30593	Human sec	934	16	61.5	312	3	AAO64668	AAO64668	Arabidops
862	16	61.5	221	4	AAO30388	AAO30388	C glutami	935	16	61.5	312	3	AAO32270	AAO32270	Arabidops
863	16	61.5	222	8	ADp31379	Adp31379	Human sec	936	16	61.5	316	3	AAO66276	AAO66276	Arabidops
864	16	61.5	222	8	ADp31300	Adp31300	Human sec	937	16	61.5	317	3	AAO64667	AAO64667	Arabidops
865	16	61.5	223	7	ABO76890	ABO76890	Pseudomon	938	16	61.5	317	3	AAO32269	AAO32269	Arabidops
866	16	61.5	226	7	ABO75147	ABO75147	Pseudomon	939	16	61.5	317	5	AAU92968	AAU92968	Arabidops
867	16	61.5	228	8	ADp31577	Adp31577	Human sec	940	16	61.5	317	7	ADp30646	ADp30646	Plant yie
868	16	61.5	228	8	ADp31334	Adp31334	Human sec	941	16	61.5	317	8	ADp41565	ADp41565	Plant tra
869	16	61.5	228	8	ADp30764	Adp30764	Human sec	942	16	61.5	317	8	ADp31615	ADp31615	Human sec
870	16	61.5	229	4	AAO91496	AAO91496	C glutami	943	16	61.5	317	8	ADp06984	ADp06984	Arabidops
871	16	61.5	229	6	ABU70511	ABU70511	Human adi	944	16	61.5	317	9	AEK26635	AEK26635	Streps to
872	16	61.5	231	5	ABJ01955	ABJ01955	158P1D7 S	945	16	61.5	318	3	AAO68087	AAO68087	Arabidops
873	16	61.5	231	8	ADp30610	Adp30610	Human sec	946	16	61.5	318	8	ADp30588	ADp30588	Human sec
874	16	61.5	233	3	AAO74791	AAO74791	Neisseria	947	16	61.5	320	8	ADp31607	ADp31607	Human sec
875	16	61.5	233	4	ADp788408	ADp788408	Novel SAR	948	16	61.5	320	8	ADp31649	ADp31649	Human sec
876	16	61.5	234	4	ABG10334	ABG10334	Novel hum	949	16	61.5	320	8	ABM82688	ABM82688	Human dia
877	16	61.5	234	8	ADp30484	Adp30484	Human sec	950	16	61.5	323	3	AAO68086	AAO68086	Arabidops
878	16	61.5	234	8	ADp31166	Adp31166	Human sec	951	16	61.5	323	4	AAAB84370	AAAB84370	Antno act
879	16	61.5	234	8	ADp31533	Adp31533	Plant ful	952	16	61.5	326	7	ABO66821	ABO66821	Klebslell
880	16	61.5	236	8	ABM84301	ABM84301	Human dia	953	16	61.5	327	8	ADp30478	ADp30478	Human sec
881	16	61.5	236	8	ABM84275	ABM84275	Human dia	954	16	61.5	327	8	ADp31660	ADp31660	Plant ful
882	16	61.5	236	8	ABM84286	ABM84286	Human dia	955	16	61.5	328	7	ABO77840	ABO77840	Pseudomon
883	16	61.5	240	7	ABO73245	ABO73245	Pseudomon	956	16	61.5	339	8	ADp30892	ADp30892	Human sec
884	16	61.5	240	8	ADp31579	Adp31579	Human sec	957	16	61.5	341	7	ABO66273	ABO66273	Klebslell
885	16	61.5	240	8	ADp31475	Adp31475	Human sec	958	16	61.5	344	3	AAO75045	AAO75045	Neisseria
886	16	61.5	243	8	ADp30983	Adp30983	Human sec	959	16	61.5	345	8	ADp31204	ADp31204	Human sec
887	16	61.5	243	8	ADp31343	Adp31343	Human sec	960	16	61.5	346	8	ADp31355	ADp31355	Human sec
888	16	61.5	246	7	ABO74848	ABO74848	Pseudomon	961	16	61.5	348	4	ABB68968	ABB68968	Drosoph11
889	16	61.5	246	7	ABO75282	ABO75282	Pseudomon	962	16	61.5	348	8	ADp34950	ADp34950	Xylanase
890	16	61.5	249	7	ABO80226	ABO80226	Pseudomon	963	16	61.5	348	8	ADp30985	ADp30985	Human sec
891	16	61.5	249	8	ADp30645	Adp30645	Human sec	964	16	61.5	348	8	ADp31676	ADp31676	Human sec
892	16	61.5	249	8	ADp30754	Adp30754	Human sec	965	16	61.5	349	8	ADp48010	ADp48010	Polypepti
893	16	61.5	252	4	ABB63479	ABB63479	Drosoph11	966	16	61.5	350	8	ADp31672	ADp31672	Human sec
894	16	61.5	253	7	ABO70290	ABO70290	Pseudomon	967	16	61.5	353	8	ADp31558	ADp31558	Human sec
895	16	61.5	254	8	ADp31396	Adp31396	Human sec	968	16	61.5	356	4	ABG04360	ABG04360	Novel hum
896	16	61.5	255	8	ADp30658	Adp30658	Human sec	969	16	61.5	358	8	ADp31063	ADp31063	Human sec
897	16	61.5	252	8	ADp31674	Adp31674	Human sec	970	16	61.5	363	7	ABO68771	ABO68771	Pseudomon
898	16	61.5	263	8	ADp31401	Adp31401	Human sec	971	16	61.5	363	8	ADp30659	ADp30659	Human sec
899	16	61.5	264	8	ADp31456	Adp31456	Human sec	972	16	61.5	366	8	ADp31106	ADp31106	Human sec
900	16	61.5	266	9	AEB37222	AEB37222	L. pneumo	973	16	61.5	366	8	ADp31199	ADp31199	Human sec

974	16	61.5	366	8	ADP31075	Adp31075 Human sec
975	16	61.5	369	7	ABO72044	AbO72044 Pseudomon
976	16	61.5	369	8	ADP31289	Adp31289 Human sec
977	16	61.5	372	8	ADP31610	Adp31610 Human sec
978	16	61.5	375	9	ADX06695	Adx06695 Cyc11n-de
979	16	61.5	380	3	AAG04919	Aag04919 Arabidops
980	16	61.5	380	3	AAG32553	Aag32553 Arabidops
981	16	61.5	381	2	AAV38570	Aav38570 Neisseria
982	16	61.5	381	2	AAV75046	Aav75046 Neisseria
983	16	61.5	381	9	AE848989	Aeb48989 N. mening
984	16	61.5	382	2	AAV38571	Aav38571 Neisseria
985	16	61.5	382	3	AAV75044	Aav75044 Neisseria
986	16	61.5	382	4	AB870181	Abb70181 Drosophil
987	16	61.5	382	9	AE848990	Aeb48990 N. gonorr
988	16	61.5	382	8	AE848990	Adp31338 Human sec
989	16	61.5	384	8	ADP31318	Adp31318 Human sec
990	16	61.5	387	8	ADP31158	Adp31158 Human sec
991	16	61.5	387	9	ADZ47819	Adz47819 Amino aci
992	16	61.5	389	8	ADX78458	Adx78458 Plant ful
993	16	61.5	389	8	ADY24278	Ady24278 Plant ful
994	16	61.5	391	7	ADN31030	Adn31030 Fruit fly
995	16	61.5	396	8	ADP31625	Adp31625 Human sec
996	16	61.5	396	8	ADP31378	Adp31378 Human sec
997	16	61.5	398	4	AAM40928	Aam40928 Human pol
998	16	61.5	398	7	ADF60215	Adf60215 Human con
999	16	61.5	398	7	ABO71161	AbO71161 Pseudomon
1000	16	61.5	398	8	ADP31604	Adp31604 Human sec

## ALIGNMENTS

RESULT 1  
ADZ47471 ID ADZ47471 standard; peptide; 30 AA.  
AC ADZ47471;  
XX  
XX 30-JUN-2005 (first entry)  
DT  
XX  
XX Alu sense peptide SEQ ID NO 50.  
DE  
XX  
XX hyaluron synthase; HAS; hyaluronidase; HYAL; HAS1; HAS2; HAS3; HYAL1;  
KM HYAL2; HAL3; PH-20; cancer; hyperproliferation; inflammation; anabolic;  
KM hypertensive; gynaecological; neuroprotective; antianemic; cycostatic;  
KM anti-inflammatory; endocrine-gen; immunosuppressive; gene therapy.  
XX  
XX Unidentified.  
OS  
XX  
XX WO2005035548-A1.  
PN  
XX  
XX 21-APR-2005.  
PD  
XX  
XX 11-OCT-2004; 2004WO-AU001383.  
PF  
XX  
XX 10-OCT-2003; 2003AU-00905551.  
PR  
XX  
XX 01-DEC-2003; 2003AU-00906658.  
PR  
XX  
XX (MEDI-) MEDITECH RES LTD.  
PA  
XX  
XX Brown TJ, Brownlee GR;  
PI  
XX  
XX WPI; 2005-315540/32.  
DR  
XX  
XX Novel compound capable of reducing level, function or activity of  
PT hyaluron synthase or hyaluronidase, useful for treating cancer,  
PT hyperproliferative condition, A-beta-lipoproteinemia, A-V, A beta-2-  
PT microglobulin amyloidosis.  
PT  
XX  
XX Example 22; SEQ ID NO 50; 277pp; English.  
PS  
XX  
XX This invention describes a novel nucleic acid molecule capable of  
CC reducing the level of hyaluron synthase (HAS) or hyaluronidase (HYAL)

or the function or activity of HAS or HYAL. HAS is selected from HAS1, HAS2 and HAS3. HYAL is chosen from HYAL1, HYAL2, HAL3 and PH-20. The nucleic acid molecule is an oligonucleotide or its chemically modified form comprising a chemically modified backbone or a non-natural internucleoside linkage. The compound is an interactive molecule capable of binding or otherwise associating with HAS and/or HYAL to reduce HAS and/or HYAL function or activity. The novel molecule is useful for treating cancer, hyperproliferative conditions or inflammatory conditions. The products of the invention are useful for treating or prophylaxis of a condition in subject, which involves administering to the subject, an HAS and/or HYAL level-reducing or HAS and/or HYAL activity reducing effective amount of active molecule. The products of the invention have anabolic, hypertensive, gynaecological, neuroprotective, antianemic, cycostatic, anti-inflammatory, endocrine-gen, and immunosuppressive activity. This sequence represents a peptide used in the method of the invention. NOTE: The specification describes SEQ ID NO 27-51 as nucleotide primers however the Sequence Listing represents the nucleotides in a three letter amino acid code. The nucleotide sequences have been made and are represented in { }.

## Sequence 30 AA;

Query Match 69.2%; Score 18; DB 9; Length 30;  
Best Local Similarity 20.0%; Pred. No. 2.4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10  
DB 18 CTAATAATAC 27

RESULT 2  
ADP30533 ID ADP30533 standard; protein; 89 AA.  
AC ADP30533;  
XX  
XX 12-AUG-2004 (first entry)  
DT  
XX  
XX Human secreted protein SEQ ID #1300.  
DE  
XX  
XX Cycostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2004035732-A2.  
PN  
XX  
XX 29-APR-2004.  
PD  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406579P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406588P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406588P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406611P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406612P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406616P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406640P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406642P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406646P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406653P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406655P.  
PR  
XX  
XX 29-AUG-2002; 2002US-0406666P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410946P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410947P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410948P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410949P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410953P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410957P.  
PR  
XX  
XX 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

XX PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI PI Halenbeck RF, Huang MM, Kochakora S, Haishan L, Linemann T;  
FI FI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX XX WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX

Claim 1; SEQ ID NO 2531; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytosratic, the  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOEMB and is not in the specification.

Sequence 89 AA;

Query Match	Similarity	69.2%	Score 18	DB 8	Length 89
Best Local	Similarity	20.0%	Pred. No. 3.1		
Matches	2	Conservative	0	Mismatches	8
				Indels	0
				Gaps	0
QY	1 CXXXXXXXXC 10				
Db	24 CTTAAAAAAC 33				
RESULT 3					
ID	ADP30531				
AC	ADP30531 standard; protein; 89 AA.				
XX	ADP30531;				
XX	12-AUG-2004 (first entry)				
XX	Human secreted protein SEQ ID #1298.				
XX	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;				
XX	cancer; inflammatory; immune; human secreted protein.				
XX	Homo sapiens.				
XX	WO2004035732-A2.				
XX	29-APR-2004.				
XX	28-AUG-2003; 2003WO-US026780.				
XX	29-AUG-2002; 2002US-0406576P.				
XX	29-AUG-2002; 2002US-0406579P.				
XX	29-AUG-2002; 2002US-0406585P.				
XX	29-AUG-2002; 2002US-0406588P.				
XX	29-AUG-2002; 2002US-0406608P.				
XX	29-AUG-2002; 2002US-0406611P.				
XX	29-AUG-2002; 2002US-0406612P.				
XX	29-AUG-2002; 2002US-0406616P.				
XX	29-AUG-2002; 2002US-0406640P.				
XX	29-AUG-2002; 2002US-0406642P.				
XX	29-AUG-2002; 2002US-0406646P.				
XX	29-AUG-2002; 2002US-0406653P.				
XX	29-AUG-2002; 2002US-0406655P.				
XX	29-AUG-2002; 2002US-0406666P.				
XX	17-SEP-2002; 2002US-0410946P.				
XX	17-SEP-2002; 2002US-0410947P.				
XX	17-SEP-2002; 2002US-0410948P.				
XX	17-SEP-2002; 2002US-0410949P.				
XX	17-SEP-2002; 2002US-0410953P.				
XX	17-SEP-2002; 2002US-0410957P.				
XX	17-SEP-2002; 2002US-0410958P.				
XX	17-SEP-2002; 2002US-0410959P.				
XX	17-SEP-2002; 2002US-0410960P.				
XX	17-SEP-2002; 2002US-0410961P.				
XX	17-SEP-2002; 2002US-0410962P.				
XX	17-SEP-2002; 2002US-0411019P.				
XX	17-SEP-2002; 2002US-0411022P.				
XX	17-SEP-2002; 2002US-0411023P.				
XX	17-SEP-2002; 2002US-0411024P.				
XX	17-SEP-2002; 2002US-0411032P.				
XX	17-SEP-2002; 2002US-0411035P.				
XX	17-SEP-2002; 2002US-0411037P.				
XX	17-SEP-2002; 2002US-0411041P.				
XX	17-SEP-2002; 2002US-0411045P.				
XX	17-SEP-2002; 2002US-0411046P.				
XX	17-SEP-2002; 2002US-0411048P.				
XX	17-SEP-2002; 2002US-0411052P.				
XX	17-SEP-2002; 2002US-0411055P.				
XX	17-SEP-2002; 2002US-0411073P.				
XX	17-SEP-2002; 2002US-0411082P.				
XX	17-SEP-2002; 2002US-0411101P.				
XX	17-SEP-2002; 2002US-0411111P.				
XX	18-APR-2003; 2003US-0463700P.				

```

PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485244P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486466P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486691P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX PI Halebbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX
XX Claim 1, SEQ ID NO 2529; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytosstatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and vincindal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 89 AA;
SQ
Query Match 69.2%; Score 18; DB 8; Length 89;
Best Local Similarity 20.0%; Pred. No. 3.1;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXC 10
Db 24 CTTAAAAAAC 33
RESULT 4
ID AAM93527 standard; protein; 109 AA.
XX
XX AAM93527;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide, SEQ ID NO: 3263.
XX
XX Human, full length cDNA; cDNA synthesis; oligo-capping.
XX
XX Homo sapiens.

```

```

XX
XX EP130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-00114089.
XX
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
XX
XX N-PSDB; AAK94457.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
XX PT in genetic manipulation.
XX
XX Claim 8; SEQ ID NO 3263; 1380bp + Sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
XX CC clones. 830 cDNA molecules encoding a human protein have been isolated
XX CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX CC been determined. Primers for synthesizing the full length cDNA are useful
XX CC for clarifying the function of the protein encoded by the cDNA. The full
XX CC length clones were obtained by construction of full length enriched cDNA
XX CC libraries that were synthesised by the oligo-capping method. The primers
XX CC enable the production of the full length cDNA easily without any special
XX CC methods. The present sequence is a polypeptide encoded by a full length
XX CC human cDNA of the invention. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in CD-ROM
XX CC format directly from EPO
XX
XX Sequence 109 AA;
SQ
Query Match 69.2%; Score 18; DB 4; Length 109;
Best Local Similarity 20.0%; Pred. No. 3.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXC 10
Db 4 CSSSARSTGC 13
RESULT 5
ID ADL31230 standard; protein; 109 AA.
XX
XX ADL31230;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human protein encoded by a full length cDNA clone Segid 3263.
XX
XX human; medicine; signal transduction; glycoprotein; transcription;
XX KW oligo-capping method.
XX
XX Homo sapiens.
XX
XX EP1396543-A2.
XX
XX 10-MAR-2004.
XX
XX 07-JUL-2000; 2003EP-00025638.
XX
XX 08-JUL-1999; 99JP-00194486.
XX 11-JAN-2000; 2000JP-00118774.
XX 02-MAY-2000; 2000JP-00183865.
XX 07-JUL-2000; 2000EP-00114089.

```



XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
 PA Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 XX WPI; 2004-204755/20.  
 DR N-PSDB; ADL31229.  
 PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full  
 length human cDNAs.  
 XX Example 1; SEQ ID NO 3263; 1340bp; English.  
 PS  
 XX This invention relates to a novel primers useful for synthesizing full  
 length cDNA molecules that encode human proteins. Specifically, it refers  
 CC to secretory or membrane proteins that are potential therapeutic agents/  
 CC target molecules in the field of medicine, and in particular genes  
 CC encoding proteins that are associated with signal transduction,  
 CC glycoproteins and transcription. The present invention describes a method  
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'  
 CC ends using the oligo-capping method. This polypeptide sequence is a full  
 length human protein of the invention.  
 CC  
 SQ Sequence 109 AA;  
 Query Match 69.2%; Score 18; DB 8; Length 109;  
 Best Local Similarity 20.0%; Pred. No. 3.2;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 Oy 1 CXXXXXXXXXC 10  
 Db 4 CSSSATSTSC 13  
 RESULT 6  
 ADP30690  
 ID ADP30690 standard; protein; 109 AA.  
 AC ADP30690;  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #1457.  
 XX  
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 cancer; inflammatory; immune; human secreted protein.  
 OS Homo sapiens.  
 PN WO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406642P.  
 PR 29-AUG-2002; 2002US-0406646P.  
 PR 29-AUG-2002; 2002US-0406653P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 29-AUG-2002; 2002US-0406666P.  
 PR 17-SEP-2002; 2002US-0410946P.  
 PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410958P.  
 PR 17-SEP-2002; 2002US-0410959P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410961P.  
 PR 17-SEP-2002; 2002US-0410962P.  
 PR 17-SEP-2002; 2002US-0411019P.  
 PR 17-SEP-2002; 2002US-0411022P.  
 PR 17-SEP-2002; 2002US-0411023P.  
 PR 17-SEP-2002; 2002US-0411024P.  
 PR 17-SEP-2002; 2002US-0411032P.  
 PR 17-SEP-2002; 2002US-0411035P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411041P.  
 PR 17-SEP-2002; 2002US-0411045P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411101P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463709P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471336P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485244P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 XX  
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX  
 PI Williams LT, Chu K, Lee E, Heslir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 DR WPI; 2004-348438/32.  
 XX  
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX  
 PS Claim 1; SEQ ID NO 2688; 428bp; English.  
 XX  
 XX The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is

CC available on WIPOMB and is not in the specification.  
XX Sequence 109 AA;  
SQ

Query Match 69.2%; Score 18; DB 8; Length 109;  
Best Local Similarity 20.0%; Pred. No. 3.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 96 CTTTATATATC 105

RESULT 7  
ID ADP31474 standard; protein; 144 AA.  
XX ADP31474;  
AC  
XX  
DT 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #2241.  
DE  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM Cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX MO2004035732-A2.  
XX  
XX  
XX 29-APR-2004.  
PD  
XX  
PF 28-AUG-2003; 2003MO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
XX Claim 1; SEQ ID NO 3472; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.

SQ Sequence 144 AA;  
XX

Query Match 69.2%; Score 18; DB 8; Length 144;  
Best Local Similarity 20.0%; Pred. No. 3.4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 72 CAATTAATC 81

RESULT 8  
ABO69151  
ID ABO69151 standard; protein; 167 AA.  
XX  
XX ABO69151;  
AC  
XX  
DT 29-JUL-2004 (first entry)  
XX  
XX Pseudomonas aeruginosa polypeptide #1326.  
DE

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
 XX Pseudomonas aeruginosa.  
 OS US6551795-B1.  
 XX  
 XX 22-APR-2003.  
 PD  
 XX 18-FEB-1999; 99US-00252991.  
 PF  
 XX 18-FEB-1998; 98US-0074788P.  
 PR 27-JUL-1998; 98US-0094190P.  
 XX  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 PA  
 XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
 PI WPI; 2003-615309/58.  
 DR N-PSDB; ABD02722.  
 DR  
 XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
 PT pathological conditions resulting from bacterial infection.  
 PT  
 XX  
 XX Disclosure; SEQ ID NO 17897; 455bp; English.  
 PS  
 XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful in diagnosis and  
 CC therapy of pathological conditions, as molecular targets for diagnostics,  
 CC prophylaxis and treatment of pathological conditions resulting from a  
 CC bacterial infection, for evaluating a compound, such as a polypeptide,  
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
 CC effective antibacterial targets, as targets for antibacterial drugs,  
 CC including anti-P. aeruginosa drugs, as templates for recombinant  
 CC production of P. aeruginosa-derived peptides or polypeptides, as target  
 CC components for diagnosis and/or treatment of P. aeruginosa-caused  
 CC infection, and in detection of P. aeruginosa sequences or other sequences  
 CC of Pseudomonas species using biotech technology. Sequences ABO67826-  
 CC ABO68396 represent P. aeruginosa polypeptides of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format from USPTO at  
 CC seqdata.uspto.gov/sequence.html  
 CC  
 XX  
 SQ Sequence 167 AA;  
 Query Match 69.2%; Score 18; DB 7; Length 167;  
 Best Local Similarity 20.0%; Pred. No. 3.5;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXX 10  
 Db 60 CISTSTTSC 69  
 Db  
 RESULT 9  
 ADS12185  
 ID ADS12185 standard; protein; 169 AA.  
 XX  
 AC ADS12185;  
 XX  
 XX 16-DEC-2004 (first entry)  
 DT  
 XX Human therapeutic contig protein - SEQ ID 2422.  
 DE  
 XX antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;  
 KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;  
 KW aplastic anaemia; cancer; wound healing; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 1. 169  
 FT

PT /label= Unknown, OTHER  
 FT /note= "OTHER = In-frame STOP codon"  
 XX  
 XX WO2004080148-A2.  
 PN  
 XX 23-SEP-2004.  
 PD  
 XX 30-SEP-2003; 2003WO-US030720.  
 PF  
 XX 02-OCT-2002; 2002US-0416186P.  
 PR  
 XX (NUVE-) NUVELO INC.  
 PA  
 XX Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y,  
 PI Wang D, Chen R, Zhao Q, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;  
 XX WPI; 2004-668857/65.  
 DR N-PSDB; ADS11587.  
 DR  
 XX  
 XX New polynucleotide, useful in preparing a composition for diagnosing or  
 PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,  
 PT aplastic anemia or cancer for promoting wound healing.  
 PT  
 XX  
 XX Example 2; SEQ ID NO 2422; 718bp; English.  
 PS  
 XX The invention relates to a novel isolated polynucleotide and the encoded  
 CC polypeptide. The molecules of the invention demonstrate antiinflammatory,  
 CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may  
 CC be useful in preparing a composition for diagnosing or treating  
 CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell  
 CC disorders, such as aplastic anaemia or cancer, as well as for promoting  
 CC wound healing. The molecules may also be utilised during gene therapy  
 CC procedures. The current sequence is that of a human therapeutic contig  
 CC protein of the invention.  
 CC  
 XX  
 SQ Sequence 169 AA;  
 Query Match 69.2%; Score 18; DB 8; Length 169;  
 Best Local Similarity 20.0%; Pred. No. 3.5;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXX 10  
 Db 81 CAAAAAACC 90  
 Db  
 RESULT 10  
 ADP30794  
 ID ADP30794 standard; protein; 171 AA.  
 XX  
 AC ADP30794;  
 XX  
 XX 12-AUG-2004 (first entry)  
 DT  
 XX Human secreted protein SEQ ID #1561.  
 DE  
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KW cancer; inflammatory; immune; human secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2004035732-A2.  
 PN  
 XX 29-APR-2004.  
 PD  
 XX 28-AUG-2003; 2003WO-US026780.  
 PF  
 XX 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406808P.  
 PR 29-AUG-2002; 2002US-0406811P.  
 PR



PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485255P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
(FIVE-) FIVE PRIME THERAPEUTICS, INC.  
XX  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halembeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX  
DR WPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 2791; 428bp; English.  
XX  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
XX  
SQ Sequence 171 AA;  
  
Query Match 69.2%; Score 18; DB 8; Length 171;  
Best Local Similarity 20.0%; Pred. No. 3.5;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

RESULT 12  
ADP30806  
ID ADP30806 standard; protein; 183 AA.  
XX  
XX  
AC ADP30806;  
XX  
XX  
DT 12-AUG-2004 (first entry)  
XX  
XX  
DE Human secreted protein SEQ ID #1573.  
XX  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO2004035732-A2.  
XX  
XX  
PD 29-APR-2004.  
XX  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410963P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 14-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 15-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LR, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2804; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.  
XX  
XX  
XX Sequence 183 AA:  
SQ  
Query Match 69.2%; Score 18; DB 8; Length 183;  
Best Local Similarity 20.0%; Pred. No. 3.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXX 10  
DB 36 CATTTC 45  
RESULT 13  
ABO76619  
ID ABO76619 standard; protein; 191 AA.  
XX  
XX ABO76619;  
AC  
XX  
XX 29-JUL-2004 (first entry)  
DE Pseudomonas aeruginosa polypeptide #8794.  
XX  
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX  
XX Pseudomonas aeruginosa.  
OS  
XX  
XX US651795-B1.  
XX  
XX 22-APR-2003.  
XX  
XX 18-FEB-1999; 99US-00252991.

PR 18-FEB-1998; 98US-0074788P.  
PR 27-JUL-1998; 98US-0094190P.  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
PI WPI; 2003-615309/58.  
DR N-Psdb; ABD10190.  
XX  
XX  
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
XX  
XX Disclosure; SEQ ID NO 25365; 455bp; English.  
XX  
XX The invention relates to Pseudomonas aeruginosa polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)  
XX  
XX  
XX Sequence 191 AA:  
SQ  
Query Match 69.2%; Score 18; DB 7; Length 191;  
Best Local Similarity 20.0%; Pred. No. 3.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXX 10  
DB 49 CSTASATSC 58  
RESULT 14  
ADZ56213  
ID ADZ56213 standard; protein; 214 AA.  
XX  
XX ADZ56213;  
AC  
XX  
XX 30-JUN-2005 (first entry)  
DE Human KIAA0779 splice variant clone CUN00149041.a, protein.  
XX  
XX KIAA0779; gene expression; inflammation; antiinflammatory; cancer;  
XX proliferation; neoplasm; cytostatic; immune disorder; immunomodulator;  
XX metabolic disorder; metabolic; viral infection; virucide; infection.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2005035569-A2.  
XX  
XX 21-APR-2005.  
XX  
XX 12-OCT-2004; 2004WO-US033408.  
XX  
XX 10-OCT-2003; 2003US-0510612P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Wong JGP, Hestir K, Collins ALT;  
PI WPI; 2005-296268/30.  
XX  
XX

DR N-PSDB; ADZ56200, ADZ56231.  
XX  
PT New isolated KIAA0779 nucleic acids and polypeptides, useful for  
PT diagnosing, preventing and/or treating inflammatory, immune, viral  
PT disorders and cancer, such as kidney, lung, ovarian, bladder, breast,  
PT prostate and skin cancers.  
XX  
PS Claim 11; SEQ ID NO 18; 121pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule comprises at  
CC least one polynucleotide sequence (appearing as ADZ56196-ADZ56197,  
CC ADZ56199-ADZ56200, ADZ56226-ADZ56227 and ADZ56230-ADZ56231), sequences  
CC hybridizing to them under high stringency conditions, sequences having at  
CC least 80% sequence identity to them, their complements or biologically  
CC active fragments. The nucleic acids are splice variants of the human  
CC KIAA0779 gene, encoding transmembrane domain protein(s). Also included  
CC are a double-stranded isolated nucleic acid molecule comprising the  
CC nucleic acid molecule cited above, a vector comprising the isolated  
CC expression of the nucleic acid molecule), a recombinant host cell  
CC comprising the nucleic acid molecule cited above, an isolated polypeptide  
CC (comprising the encoded proteins or the non-transmembrane (TM) regions of  
CC the proteins), a method of making a recombinant host cell, a method of  
CC making a polypeptide, a method of determining the presence of the nucleic  
CC acid molecule cited above in a sample, a method of determining the  
CC presence of a specific antibody to the polypeptide of (4) in a sample, a  
CC method of determining the presence of the polypeptide in a sample, an  
CC antibody specifically binding to and/or interfering with the biological  
CC activity of the nucleic acid molecule cited above (or the polypeptide or  
CC its biologically active fragment), a composition comprising a  
CC pharmaceutical carrier or excipient (and one or more active agents chosen  
CC from the nucleic acid molecule cited above, the vector, the polypeptide,  
CC and the antibody), a bacteriophage comprising the antibody (or its  
CC fragment), a bacterial cell comprising the bacteriophage, a recombinant  
CC host cell that produces the antibody, an animal injected with one or more  
CC active agents (chosen from the nucleic acid molecule, the vector, the  
CC host cell, the polypeptide, and the antibody), a diagnostic kit  
CC (comprising a nucleic acid molecule having at least 6 contiguous  
CC nucleotides from the nucleic acid molecule cited above, the isolated  
CC polypeptide, or antibody, and reagents to carry out an immunoassay), a  
CC method of making an antibody, a method of identifying a modulating agent  
CC that modulates the biological activity of the polypeptide, a modulator  
CC composition comprising a modulator and a pharmaceutical carrier, a method  
CC of treating a disease in a subject, and a method of treating cancer,  
CC preferably kidney, cervical, squamous lung, ovarian, bladder, breast,  
CC endometrial, prostate, and skin cancer, in a subject. The methods and  
CC compositions of the present invention are useful for the diagnosis,  
CC prevention and/or treatment of inflammatory, immune, viral disorders and  
CC cancer, such as kidney, cervical, squamous lung, ovarian, bladder,  
CC breast, endometrial, prostate, and skin cancer. The present sequence  
CC represents a protein encoded by a cDNA splice variant from the KIAA0779  
CC gene.  
XX  
SQ Sequence 214 AA;  
XX  
Query Match 69.2%; Score 18; DB 9; Length 214;  
Best Local Similarity 20.0%; Pred. No. 3.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXX 10  
Db 173 CAAAAAAC 182  
RESULT 15  
ADP30921  
ID ADP30921 standard; protein; 228 AA.  
XX  
AC ADP30921;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SBO ID #1688.

XX  
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Viaricide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406589P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 22-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485225P.  
PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,  
XX  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2919; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.  
XX  
SQ Sequence 228 AA;  
  
Query Match 69.24; Score 18; DB 8; Length 228;  
Best Local Similarity 20.0%; Pred. No. 3.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXX 10  
Db 54 CATMAATTAC 63  
  
RESULT 16  
ADP31485  
ID ADP31485 standard; protein; 252 AA.  
XX  
AC ADP31485;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2252.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410966P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471106P.  
PR 19-MAY-2003; 2003US-0471136P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485234P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
XX Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
XX Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3483; 428bp; English.  
XX



CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMB and is not in the specification.  
 XX  
 SQ Sequence 252 AA;  
 Query Match 69.2%; Score 18; DB 8; Length 252;  
 Best Local Similarity 20.0%; Pred. No. 3.9;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXXC 10  
 Db 210 CTTTATATAC 219  
 RESULT 17  
 ADP30479  
 ID ADP30479 standard; protein; 258 AA.  
 AC ADP30479;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #1246.  
 XX  
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KW cancer; inflammatory; immune; human secreted protein.  
 OS Homo sapiens.  
 XX  
 PN WO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 XX 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406642P.  
 PR 29-AUG-2002; 2002US-0406646P.  
 PR 29-AUG-2002; 2002US-0406653P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 29-AUG-2002; 2002US-0406666P.  
 PR 17-SEP-2002; 2002US-0410946P.  
 PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410948P.  
 PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410958P.  
 PR 17-SEP-2002; 2002US-0410959P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410961P.  
 PR 17-SEP-2002; 2002US-0410962P.  
 PR 17-SEP-2002; 2002US-0411019P.  
 PR 17-SEP-2002; 2002US-0411022P.  
 PR 17-SEP-2002; 2002US-0411023P.  
 PR 17-SEP-2002; 2002US-0411024P.  
 PR 17-SEP-2002; 2002US-0411032P.  
 PR 17-SEP-2002; 2002US-0411035P.  
 PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.  
 PR 17-SEP-2002; 2002US-0411045P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411101P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-047136P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485252P.  
 PR 14-JUL-2003; 2003US-0486466P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 XX  
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX  
 PA William LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Huang MM, Kochakota S, Hsiehan L, Linnemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 DR WPI; 2004-348438/32.  
 XX  
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX  
 PS Claim 1; SEQ ID NO 2477; 428bp; English.  
 XX  
 CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMB and is not in the specification.  
 XX  
 SQ Sequence 258 AA;  
 Query Match 69.2%; Score 18; DB 8; Length 258;  
 Best Local Similarity 20.0%; Pred. No. 3.9;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXXC 10  
 Db 237 CTTTATATAC 246  
 RESULT 18  
 ADP31412

ID ADP31412 standard; protein; 264 AA.  
 XX  
 AC ADP31412;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #2179.  
 XX  
 KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 cancer; inflammatory; immune; human secreted protein.  
 OS Homo sapiens.  
 XX  
 PN WO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 XX 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406642P.  
 PR 29-AUG-2002; 2002US-0406646P.  
 PR 29-AUG-2002; 2002US-0406633P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 29-AUG-2002; 2002US-0406666P.  
 PR 17-SEP-2002; 2002US-0410946P.  
 PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410948P.  
 PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410958P.  
 PR 17-SEP-2002; 2002US-0410959P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410961P.  
 PR 17-SEP-2002; 2002US-0410962P.  
 PR 17-SEP-2002; 2002US-0411019P.  
 PR 17-SEP-2002; 2002US-0411022P.  
 PR 17-SEP-2002; 2002US-0411023P.  
 PR 17-SEP-2002; 2002US-0411024P.  
 PR 17-SEP-2002; 2002US-0411032P.  
 PR 17-SEP-2002; 2002US-0411035P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411041P.  
 PR 17-SEP-2002; 2002US-0411045P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411101P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463733P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471336P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486893P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 XX  
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX  
 XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
 PI Halenbeck RF, Huang MM, Kothakota S, Hsiehan L, Linnemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 XX  
 XX WPI; 2004-348438/32.  
 DR  
 XX  
 XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX  
 XX Claim 1; SEQ ID NO 3410; 428bp; English.  
 PS  
 XX  
 CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMEB and is not in the specification.  
 XX  
 XX Sequence 264 AA;  
 SQ  
 Query Match 69.2%; Score 18; DB 8; Length 264;  
 Best Local Similarity 20.0%; Pred. No. 3.9;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXX 10  
 DB 186 CATTAATTC 195  
 RESULT 19  
 ADP30500  
 ID ADP30500 standard; protein; 270 AA.  
 XX  
 AC ADP30500;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #1267.  
 XX  
 KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 cancer; inflammatory; immune; human secreted protein.  
 OS Homo sapiens.  
 XX  
 PN WO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 XX 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485252P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,  
DR WPI; 2004-348438/32.

XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2498; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMED and is not in the specification.  
XX  
SQ Sequence 270 AA;  
Query Match 69.2%; Score 18; DB 8; Length 270;  
Best Local Similarity 20.0%; Pred. No. 3.9;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXC 10  
Db 233 CTTATATATC 242  
RESULT 20  
ID ADP31321 standard; protein; 270 AA.  
XX  
XX ADP31321;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #2088.  
DE  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX MO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003MO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486981P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MW, Kochakota S, Haishan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI, 2004-348438/32.  
XX  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3319; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOBBS and is not in the specification.  
XX  
XX Sequence 270 AA:

Query Match 69.2%; Score 18; DB 8; Length 270;  
Best Local Similarity 20.0%; Pred. No. 3.9;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10  
Db 245 CTTTATATAC 254  
  
RESULT 21  
ADP31473  
ID ADP31473 standard; protein; 294 AA.  
XX  
AC ADP31473;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2240.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN MO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 09-JUN-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 08-JUL-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MW, Kochakota S, Halshan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

PS Claim 1; SEQ ID NO 3471; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.

XX Sequence 294 AA;

Query Match 69.2%; Score 18; DB 8; Length 294;

Best Local Similarity 20.0%; Pred. No. 4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 10

DB 72 CATTAATC 81

RESULT 22

ID ADP31192 standard; protein; 297 AA.

AC ADP31192;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1959.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW Cancer; inflammatory; immune; human secreted protein.

XX Homo sapiens.

XX WO2004035732-A2.

XX

PD 29-APR-2004.

XX 28-AUG-2003; 2003MO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

XX 29-AUG-2002; 2002US-0406579P.

XX 29-AUG-2002; 2002US-0406585P.

XX 29-AUG-2002; 2002US-0406588P.

XX 29-AUG-2002; 2002US-0406608P.

XX 29-AUG-2002; 2002US-0406611P.

XX 29-AUG-2002; 2002US-0406612P.

XX 29-AUG-2002; 2002US-0406616P.

XX 29-AUG-2002; 2002US-0406640P.

XX 29-AUG-2002; 2002US-0406642P.

XX 29-AUG-2002; 2002US-0406646P.

XX 29-AUG-2002; 2002US-0406653P.

XX 29-AUG-2002; 2002US-0406655P.

XX 29-AUG-2002; 2002US-0406666P.

XX 17-SEP-2002; 2002US-0410946P.

XX 17-SEP-2002; 2002US-0410947P.

XX 17-SEP-2002; 2002US-0410948P.

XX 17-SEP-2002; 2002US-0410949P.

XX 17-SEP-2002; 2002US-0410953P.

XX 17-SEP-2002; 2002US-0410957P.

XX 17-SEP-2002; 2002US-0410958P.

XX 17-SEP-2002; 2002US-0410959P.

XX 17-SEP-2002; 2002US-0410960P.

XX 17-SEP-2002; 2002US-0410961P.

XX 17-SEP-2002; 2002US-0410962P.

XX 17-SEP-2002; 2002US-0411019P.

XX 17-SEP-2002; 2002US-0411022P.

XX 17-SEP-2002; 2002US-0411023P.

XX 17-SEP-2002; 2002US-0411024P.

XX 17-SEP-2002; 2002US-0411032P.

XX 17-SEP-2002; 2002US-0411035P.

XX 17-SEP-2002; 2002US-0411037P.

XX 17-SEP-2002; 2002US-0411041P.

XX 17-SEP-2002; 2002US-0411045P.

XX 17-SEP-2002; 2002US-0411046P.

XX 17-SEP-2002; 2002US-0411048P.

XX 17-SEP-2002; 2002US-0411052P.

XX 17-SEP-2002; 2002US-0411052P.

XX 17-SEP-2002; 2002US-0411055P.

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halebek R, Huang MM, Kochakota S, Haisan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
PS Claim 1; SEQ ID NO 3190; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMER and is not in the specification.  
XX  
SQ Sequence 297 AA;  
Query Match 69.2%; Score 18; DB 8; Length 297;  
Best Local Similarity 20.0%; Pred. No. 4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXC 10  
DB 166 CTTATTATTC 175  
DE Human protein sequence SEQ ID NO:800.  
RESULT 23  
AAM25285  
ID AAM25285 standard; protein: 328 AA.  
AC AAM25285;  
XX  
XX 16-OCT-2001 (first entry)  
DT  
XX  
DE Human protein sequence SEQ ID NO:800.  
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
KW antibacterial; endocrine; cardiac; central nervous system; vitricide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
KW antiaggregant; haemostatic; vulnary; antileuk; osteopathic; eczema;  
KW dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;  
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
KW immunostimulant; gene therapy; antileuk therapy; vaccine; inflammation;  
KW anaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200153455-A2.  
PN  
XX  
XX 26-JUL-2001.  
PD  
XX  
XX 22-DEC-2000; 2000WO-US035017.  
PF  
XX  
XX 23-DEC-1999; 99US-00471275.  
PR  
XX 21-JAN-2000; 2000US-0048725.  
PR  
XX 25-APR-2000; 2000US-00552317.  
XX  
XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;  
PI WPI; 2001-457603/49.  
DR N-PSDB; AAM99226.  
XX  
XX Isolated human polynucleotides encoding polypeptides, useful for the  
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.  
PT Claim 20; Page 187; 1217bp; English.  
PS  
XX  
XX AAM99166 to AAM99904 encode the human proteins given in AAM25225 to  
CC AAM25963. The proteins can have activities based on the tissues and cells  
CC they are expressed in, such as: antiinflammatory; antirheumatic;  
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;  
CC central nervous system; vitricide; anti-HIV; fungicide; antimutagen;  
CC cardiovascular; antianaemic; antileuk; antiallergic; antiaesthetic;  
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
CC encoding them can be used in gene therapy, antileuk therapy and vaccine  
CC production. The proteins and polynucleotides are useful for screening for  
CC agonists or antagonists of a protein and for the treatment and diagnosis  
CC of disorders associated with the activity of a protein e.g. inflammation,  
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,  
CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders  
XX  
SQ Sequence 328 AA;  
Query Match 69.2%; Score 18; DB 4; Length 328;  
Best Local Similarity 20.0%; Pred. No. 4.1;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXC 10  
DB 63 CAAAAAAAC 72  
DE Human protein sequence SEQ ID NO:800.  
RESULT 24  
ADP30702  
ID ADP30702 standard; protein: 339 AA.  
AC ADP30702;  
XX  
XX 12-AUG-2004 (first entry)  
DT  
XX  
DE Human secreted protein SEQ ID #1469.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Vitricide;  
KW cancer; inflammatory; immune; human secreted protein.  
KW  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2004035732-A2.  
PN  
XX  
XX 29-APR-2004.  
PD  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR  
XX 29-AUG-2002; 2002US-0406579P.  
PR  
XX 29-AUG-2002; 2002US-0406585P.  
PR  
XX 29-AUG-2002; 2002US-0406588P.  
PR  
XX 29-AUG-2002; 2002US-0406608P.  
PR  
XX 29-AUG-2002; 2002US-0406611P.  
PR  
XX 29-AUG-2002; 2002US-0406612P.  
PR  
XX 29-AUG-2002; 2002US-0406616P.  
PR

PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halebek R, Huang MM, Kochakota S, Haishan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX MPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX

PS Claim 1; SEQ ID NO 2700; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antitumorigenic, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 339 AA;  
Query Match 69.2%; Score 18; DB 8; Length 339;  
Best Local Similarity 20.0%; Pred. No. 4.1;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXX 10  
Db 89 CAATTAATTC 98  
RESULT 25  
ID ADP31441 standard; protein; 348 AA.  
XX  
AC ADP31441;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2208.  
XX  
KW Cytostatic; Antitumorigenic; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN MO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003MO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406589P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406659P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411035P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411041P.  
 PR 17-SEP-2002; 2002US-0411045P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411081P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-047136P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486911P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 XX  
 XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX  
 XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T;  
 PI Piere K, Wang Y, Wong JGP, Wu G, Zhang H;  
 XX  
 DR WPI; 2004-348438/32.  
 XX  
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX  
 PS Claim 1; SEQ ID NO 3439; 428bp; English.  
 XX  
 CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytotoxic,  
 CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMB and is not in the specification.  
 XX  
 SQ Sequence 348 AA;  
 Query Match 69.2%; Score 18; DB 8; Length 348;  
 Best Local Similarity 20.0%; Pred. No. 4.1;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXX 10  
 DB 25 CTTAATTAAC 34

RESULT 26  
 ABG21039  
 ID ABG21039 standard; protein; 356 AA.  
 XX  
 AC ABG21039;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #21030.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 ER 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS85226.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID NO 51398; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
 CC patent did not appear in the invention. Note: The sequence data for this  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pat\_sequences  
 XX  
 SQ Sequence 356 AA;  
 Query Match 69.2%; Score 18; DB 4; Length 356;  
 Best Local Similarity 20.0%; Pred. No. 4.2;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXX 10  
 DB 183 CAAAAAASC 192  
 RESULT 27  
 ADP31267



ID ADP31267 standard; protein; 357 AA.  
 XX  
 AC ADP31267;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #2034.  
 XX  
 KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 cancer; inflammatory; immune; human secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.  
 PR 29-AUG-2002; 2002US-0406588P.  
 PR 29-AUG-2002; 2002US-0406608P.  
 PR 29-AUG-2002; 2002US-0406611P.  
 PR 29-AUG-2002; 2002US-0406612P.  
 PR 29-AUG-2002; 2002US-0406616P.  
 PR 29-AUG-2002; 2002US-0406640P.  
 PR 29-AUG-2002; 2002US-0406642P.  
 PR 29-AUG-2002; 2002US-0406646P.  
 PR 29-AUG-2002; 2002US-0406653P.  
 PR 29-AUG-2002; 2002US-0406655P.  
 PR 29-AUG-2002; 2002US-0406666P.  
 PR 17-SEP-2002; 2002US-0410946P.  
 PR 17-SEP-2002; 2002US-0410947P.  
 PR 17-SEP-2002; 2002US-0410948P.  
 PR 17-SEP-2002; 2002US-0410949P.  
 PR 17-SEP-2002; 2002US-0410953P.  
 PR 17-SEP-2002; 2002US-0410957P.  
 PR 17-SEP-2002; 2002US-0410958P.  
 PR 17-SEP-2002; 2002US-0410959P.  
 PR 17-SEP-2002; 2002US-0410960P.  
 PR 17-SEP-2002; 2002US-0410961P.  
 PR 17-SEP-2002; 2002US-0410962P.  
 PR 17-SEP-2002; 2002US-0411035P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411041P.  
 PR 17-SEP-2002; 2002US-0411045P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411101P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-047136P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 14-JUL-2003; 2003US-0486486P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 XX  
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 XX  
 PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
 PI Halenbeck RF, Huang MM, Kothakota S, Haisan L, Linnemann T,  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 XX  
 DR MPI; 2004-348438/32.  
 XX  
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX  
 PS Claim 1; SEQ ID NO 3265; 428bp; English.  
 XX  
 CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMEB and is not in the specification.  
 XX  
 SQ Sequence 357 AA;  
 QY 1 CXXXXXXXXX 69.2%; Score 18; DB 8; Length 357;  
 Db 14 CATATATATAC 23 Best Local Similarity 20.0%; Pred. No. 4.2;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 RESULT 28  
 ID ADP30505 standard; protein; 357 AA.  
 XX  
 AC ADP30505;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #1272.  
 XX  
 KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 cancer; inflammatory; immune; human secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004035732-A2.  
 XX  
 PD 29-APR-2004.  
 XX  
 PF 28-AUG-2003; 2003WO-US026780.  
 XX  
 PR 29-AUG-2002; 2002US-0406576P.  
 PR 29-AUG-2002; 2002US-0406579P.  
 PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476611P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486691P.  
PR 15-JUL-2003; 2003US-0486696P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang WM, Kotlakota S, Haishan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H,  
XX  
XX WPI; 2004-348438/32.

XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2503; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMER and is not in the specification.  
XX  
SQ Sequence 357 AA;  
XX  
Query Match 69.2%; Score 18; DB 8; Length 357;  
Best Local Similarity 20.0%; Pred. No. 4.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXC 10  
Db 169 CTAATTTC 178  
RESULT 29  
ADP31439  
ID ADP31439 standard; protein; 360 AA.  
XX  
AC ADP31439;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2206.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
PN W02004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.  
 PR 17-SEP-2002; 2002US-0411022P.  
 PR 17-SEP-2002; 2002US-0411023P.  
 PR 17-SEP-2002; 2002US-0411024P.  
 PR 17-SEP-2002; 2002US-0411032P.  
 PR 17-SEP-2002; 2002US-0411035P.  
 PR 17-SEP-2002; 2002US-0411037P.  
 PR 17-SEP-2002; 2002US-0411041P.  
 PR 17-SEP-2002; 2002US-0411045P.  
 PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411010P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471366P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.  
 (FIVE-) FIVE PRIME THERAPEUTICS INC.  
 PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
 PI Halenbeck RF, Huang MM, Kochakota S, Haisnan L, Linnemann T,  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 XX WPI; 2004-348438/32.  
 DR  
 XX  
 PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.  
 XX  
 PS Claim 1; SEQ ID NO 3437; 428bp; English.  
 XX  
 CC The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMEB and is not in the specification.  
 XX  
 SQ Sequence 360 AA;  
 Query Match 69.2%; Score 18; DB 8; Length 360;  
 Best Local Similarity 20.0%; Prod. No. 4.2;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10  
 Db 70 CTATTAAAC 79  
 RESULT 30  
 ADP31218  
 ID ADP31218 standard; protein; 390 AA.  
 XX  
 AC ADP31218;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Human secreted protein SEQ ID #1985.  
 XX  
 KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 XX Cancer; Inflammatory; Immune; human secreted protein.  
 OS Homo sapiens.  
 XX  
 PN MO2004035732-AZ.  
 PD 29-APR-2004.  
 XX  
 XX 28-AUG-2003; 2003US-05026780.  
 PF  
 XX 29-AUG-2002; 2002US-0406576P.  
 XX 29-AUG-2002; 2002US-0406579P.  
 XX 29-AUG-2002; 2002US-0406585P.  
 XX 29-AUG-2002; 2002US-0406588P.  
 XX 29-AUG-2002; 2002US-0406611P.  
 XX 29-AUG-2002; 2002US-0406612P.  
 XX 29-AUG-2002; 2002US-0406612P.  
 XX 29-AUG-2002; 2002US-0406640P.  
 XX 29-AUG-2002; 2002US-0406642P.  
 XX 29-AUG-2002; 2002US-0406646P.  
 XX 29-AUG-2002; 2002US-0406653P.  
 XX 29-AUG-2002; 2002US-0406655P.  
 XX 29-AUG-2002; 2002US-0406666P.  
 XX 17-SEP-2002; 2002US-0410946P.  
 XX 17-SEP-2002; 2002US-0410947P.  
 XX 17-SEP-2002; 2002US-0410948P.  
 XX 17-SEP-2002; 2002US-0410949P.  
 XX 17-SEP-2002; 2002US-0410953P.  
 XX 17-SEP-2002; 2002US-0410957P.  
 XX 17-SEP-2002; 2002US-0410958P.  
 XX 17-SEP-2002; 2002US-0410959P.  
 XX 17-SEP-2002; 2002US-0410960P.  
 XX 17-SEP-2002; 2002US-0410961P.  
 XX 17-SEP-2002; 2002US-0410962P.  
 XX 17-SEP-2002; 2002US-0411019P.  
 XX 17-SEP-2002; 2002US-0411022P.  
 XX 17-SEP-2002; 2002US-0411023P.  
 XX 17-SEP-2002; 2002US-0411024P.  
 XX 17-SEP-2002; 2002US-0411032P.  
 XX 17-SEP-2002; 2002US-0411033P.  
 XX 17-SEP-2002; 2002US-0411037P.  
 XX 17-SEP-2002; 2002US-0411041P.  
 XX 17-SEP-2002; 2002US-0411045P.  
 XX 17-SEP-2002; 2002US-0411046P.  
 XX 17-SEP-2002; 2002US-0411048P.  
 XX 17-SEP-2002; 2002US-0411052P.  
 XX 17-SEP-2002; 2002US-0411055P.  
 XX 17-SEP-2002; 2002US-0411073P.  
 XX 17-SEP-2002; 2002US-0411082P.  
 XX 17-SEP-2002; 2002US-0411010P.  
 XX 17-SEP-2002; 2002US-0411111P.  
 XX 18-APR-2003; 2003US-0463700P.  
 XX 18-APR-2003; 2003US-0463708P.  
 XX 18-APR-2003; 2003US-0463716P.  
 XX 18-APR-2003; 2003US-0463732P.  
 XX 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485225P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3216; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
XX antitumorigenic, immunosuppressive, antibacterial and antiviral. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present  
XX sequence represents a human secreted protein. The present sequence is  
XX available on WIPOMB and is not in the specification.  
SQ Sequence 390 AA;  
Query Match 69.2%; Score 18; DB 8; Length 390;  
Best Local Similarity 20.0%; Pred. No. 4.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXX 10  
Db 60 CAAAAAATTC 69  
RESULT 31  
ID ADP31159 standard; protein; 421 AA.  
XX  
XX ADP31159;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #1926.  
XX  
XX Cytostatic; Antitumorigenic; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX  
XX OS  
XX PN MO2004035732-A2.  
XX

PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411101P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 18-APR-2003; 2003US-0463700P.  
XX 18-APR-2003; 2003US-0463708P.  
XX 18-APR-2003; 2003US-0463716P.  
XX 18-APR-2003; 2003US-0463732P.  
XX 02-MAY-2003; 2003US-0467199P.  
XX 02-MAY-2003; 2003US-0467201P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 02-MAY-2003; 2003US-0467230P.  
XX 02-MAY-2003; 2003US-0467230P.  
XX 19-MAY-2003; 2003US-0471306P.  
XX 19-MAY-2003; 2003US-0471366P.  
XX 22-MAY-2003; 2003US-0472420P.  
XX 22-MAY-2003; 2003US-0472430P.  
XX 09-JUN-2003; 2003US-0476609P.  
XX 09-JUN-2003; 2003US-0476641P.  
XX 08-JUL-2003; 2003US-0485218P.  
XX 08-JUL-2003; 2003US-0485223P.  
XX 08-JUL-2003; 2003US-0485224P.  
XX 08-JUL-2003; 2003US-0485225P.  
XX 14-JUL-2003; 2003US-0486446P.  
XX 14-JUL-2003; 2003US-0486480P.  
XX 15-JUL-2003; 2003US-0486891P.  
XX 15-JUL-2003; 2003US-0486960P.  
XX 08-AUG-2003; 2003US-0493341P.  
XX 08-AUG-2003; 2003US-0493370P.  
XX 08-AUG-2003; 2003US-0493573P.  
XX 08-AUG-2003; 2003US-0493577P.  
XX



XX SQ Sequence 426 AA;  
Query Match 69.2%; Score 18; DB 8; Length 426;  
Best Local Similarity 20.0%; Pred. No. 4.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXX 10  
DB 66 CATTAAATTC 75  
RESULT 33  
ABO82561  
ID ABO82561 standard; protein; 438 AA.  
XX AC ABO82561;  
XX DT 29-JUL-2004 (first entry)  
XX DE Pseudomonas aeruginosa polypeptide #14736.  
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX OS Pseudomonas aeruginosa.  
XX PN US6551795-B1.  
XX PD 22-APR-2003.  
XX PF 18-FEB-1999; 99US-00252991.  
XX PR 18-FEB-1998; 98US-0074788P.  
XX PR 27-JUL-1998; 98US-0094190P.  
XX PA (GENO-) GENOME THERAPEUTICS CORP.  
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX DR WPI, 2003-615309/58.  
XX DR N-PSDB; ABD16132.  
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX PT pathological conditions resulting from bacterial infection.  
XX PS Disclosure; SEQ ID NO 31307; 455pp; English.  
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and  
XX CC therapy of pathological conditions, as molecular targets for diagnostics,  
XX CC prophylaxis and treatment of pathological conditions resulting from a  
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,  
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX CC effective antibacterial targets, as targets for antibacterial drugs,  
XX CC including anti-P. aeruginosa drugs, as templates for recombinant  
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target  
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused  
XX CC infection, and in detection of P. aeruginosa sequences or other sequences  
XX CC of Pseudomonas species using biotech technology. Sequences ABO67826-  
XX CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX CC sequence data for this patent did not form part of the printed  
XX CC specification but was obtained in electronic format from USPTO at  
XX CC seqdata.uspto.gov/sequence.html  
XX SQ Sequence 438 AA;  
Query Match 69.2%; Score 18; DB 7; Length 438;  
Best Local Similarity 20.0%; Pred. No. 4.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXX 10

DB 56 CSAATASRC 65  
RESULT 34  
ADP31465  
ID ADP31465 standard; protein; 453 AA.  
XX AC ADP31465;  
XX DT 12-AUG-2004 (first entry)  
XX DE Human secreted protein SEQ ID #2232.  
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX KW cancer; inflammatory; immune; human secreted protein.  
XX OS Homo sapiens.  
XX PN WO200403732-A2.  
XX PD 29-APR-2004.  
XX PF 28-AUG-2003; 2003WO-US026780.  
XX PR 29-AUG-2002; 2002US-0406576P.  
XX PR 29-AUG-2002; 2002US-0406579P.  
XX PR 29-AUG-2002; 2002US-0406585P.  
XX PR 29-AUG-2002; 2002US-0406588P.  
XX PR 29-AUG-2002; 2002US-0406608P.  
XX PR 29-AUG-2002; 2002US-0406611P.  
XX PR 29-AUG-2002; 2002US-0406612P.  
XX PR 29-AUG-2002; 2002US-0406616P.  
XX PR 29-AUG-2002; 2002US-0406642P.  
XX PR 29-AUG-2002; 2002US-0406646P.  
XX PR 29-AUG-2002; 2002US-0406653P.  
XX PR 29-AUG-2002; 2002US-0406655P.  
XX PR 29-AUG-2002; 2002US-0410946P.  
XX PR 17-SEP-2002; 2002US-0410947P.  
XX PR 17-SEP-2002; 2002US-0410948P.  
XX PR 17-SEP-2002; 2002US-0410949P.  
XX PR 17-SEP-2002; 2002US-0410953P.  
XX PR 17-SEP-2002; 2002US-0410957P.  
XX PR 17-SEP-2002; 2002US-0410958P.  
XX PR 17-SEP-2002; 2002US-0410959P.  
XX PR 17-SEP-2002; 2002US-0410960P.  
XX PR 17-SEP-2002; 2002US-0410961P.  
XX PR 17-SEP-2002; 2002US-0410962P.  
XX PR 17-SEP-2002; 2002US-0411019P.  
XX PR 17-SEP-2002; 2002US-0411022P.  
XX PR 17-SEP-2002; 2002US-0411024P.  
XX PR 17-SEP-2002; 2002US-0411032P.  
XX PR 17-SEP-2002; 2002US-0411035P.  
XX PR 17-SEP-2002; 2002US-0411037P.  
XX PR 17-SEP-2002; 2002US-0411041P.  
XX PR 17-SEP-2002; 2002US-0411045P.  
XX PR 17-SEP-2002; 2002US-0411046P.  
XX PR 17-SEP-2002; 2002US-0411048P.  
XX PR 17-SEP-2002; 2002US-0411052P.  
XX PR 17-SEP-2002; 2002US-0411055P.  
XX PR 17-SEP-2002; 2002US-0411073P.  
XX PR 17-SEP-2002; 2002US-0411082P.  
XX PR 17-SEP-2002; 2002US-0411101P.  
XX PR 17-SEP-2002; 2002US-0411111P.  
XX PR 18-APR-2003; 2003US-0463700P.  
XX PR 18-APR-2003; 2003US-0463708P.  
XX PR 18-APR-2003; 2003US-0463716P.  
XX PR 18-APR-2003; 2003US-0463732P.  
XX PR 02-MAY-2003; 2003US-0467199P.  
XX PR 02-MAY-2003; 2003US-0467201P.  
XX PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476509P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang WM, Kochakota S, Haislan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX MPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3463; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX CC encoding a polypeptide which is believed to be cytostatic,  
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX CC composition and methods are useful for diagnosing, preventing and  
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present  
XX CC sequence represents a human secreted protein. The present sequence is  
XX CC available on WIPWEB and is not in the specification.  
SQ Sequence 453 AA;  
Query Match 69.2%; Score 18; DB 8; Length 453;  
Best Local Similarity 20.0%; Pred. No. 4.4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXC 10  
DB 312 CAAATTATTC 321  
RESULT 35  
ADP30854  
ID ADP30854 standard; protein; 471 AA.  
XX  
AC ADP30854;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1621.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
XX MO2004035732-A2.  
PN  
PD 29-APR-2004.  
XX

PF 28-AUG-2003; 2003MO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411033P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411101P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 18-APR-2003; 2003US-0463700P.  
XX 18-APR-2003; 2003US-0463708P.  
XX 18-APR-2003; 2003US-0463716P.  
XX 18-APR-2003; 2003US-0463732P.  
XX 02-MAY-2003; 2003US-0467199P.  
XX 02-MAY-2003; 2003US-0467201P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 02-MAY-2003; 2003US-0467230P.  
XX 19-MAY-2003; 2003US-0471306P.  
XX 19-MAY-2003; 2003US-0471336P.  
XX 22-MAY-2003; 2003US-0472420P.  
XX 22-MAY-2003; 2003US-0472430P.  
XX 09-JUN-2003; 2003US-0476509P.  
XX 09-JUN-2003; 2003US-0476641P.  
XX 08-JUL-2003; 2003US-0485218P.  
XX 08-JUL-2003; 2003US-0485223P.  
XX 08-JUL-2003; 2003US-0485224P.  
XX 08-JUL-2003; 2003US-0485325P.  
XX 14-JUL-2003; 2003US-0486446P.  
XX 14-JUL-2003; 2003US-0486480P.  
XX 15-JUL-2003; 2003US-0486891P.  
XX 15-JUL-2003; 2003US-0486960P.  
XX 08-AUG-2003; 2003US-0493341P.  
XX 08-AUG-2003; 2003US-0493370P.  
XX 08-AUG-2003; 2003US-0493573P.  
XX 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX

Pt	Williams JT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
Pt	Halenbeck RF, Huang MM, Kothakota S, Haishan L, Linemann T,
Pt	Pierce K, Wang Y, Wong UGP, Wu G, Zhang H;
DR	WPI; 2004-348438/32.
XX	
Pt	New nucleic acid molecule for diagnosing, preventing or treating diseases
Pt	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
Pt	genetic, bacterial and viral disease.
XX	
Pt	
Pt	Claim 1; SEQ ID NO 2852; 428bp; English.
XX	
CC	The present invention relates to an isolated nucleic acid molecule
CC	encoding a polypeptide which is believed to be cytostatic,
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC	composition and methods are useful for diagnosing, preventing and
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,
CC	immune, metabolic, genetic, bacterial and viral diseases. The present
CC	sequence represents a human secreted protein. The present sequence is
CC	avalable on WIPWEB and is not in the specification.
XX	
SQ	Sequence 471 AA;
Query Match	69.2%; Score 18; DB 8; Length 471;
Best Local Similarity	20.0%; Pred. No. 4.4;
Matches	2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy	1 CXXXXXXXKC 10
Dd	142 CAATATATTC 151
RESULT 36	
ID	ADP31227 standard; protein; 525 AA.
XX	
AC	ADP31227;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Human secreted protein SEQ ID #1994.
XX	
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX	cancer; Inflammatory; Immune; human secreted protein.
XX	
OS	Homo sapiens.
PN	WO2004035732-A2.
XX	
PD	29-APR-2004.
XX	
PF	28-AUG-2003; 2003WO-US026780.
XX	
PR	29-AUG-2002; 2002US-0406576P.
PR	29-AUG-2002; 2002US-0406579P.
PR	29-AUG-2002; 2002US-0406585P.
PR	29-AUG-2002; 2002US-0406588P.
PR	29-AUG-2002; 2002US-0406608P.
PR	29-AUG-2002; 2002US-0406611P.
PR	29-AUG-2002; 2002US-0406612P.
PR	29-AUG-2002; 2002US-0406616P.
PR	29-AUG-2002; 2002US-0406640P.
PR	29-AUG-2002; 2002US-0406642P.
PR	29-AUG-2002; 2002US-0406646P.
PR	29-AUG-2002; 2002US-0406653P.
PR	29-AUG-2002; 2002US-0406655P.
PR	29-AUG-2002; 2002US-0406666P.
PR	17-SEP-2002; 2002US-0410946P.
PR	17-SEP-2002; 2002US-0410947P.
PR	17-SEP-2002; 2002US-0410948P.
PR	17-SEP-2002; 2002US-0410949P.
PR	17-SEP-2002; 2002US-0410953P.
PR	17-SEP-2002; 2002US-0410957P.

Sequence 525 AA:

(FIVE-) FIVE PRIME THERAPEUTICS INC.

Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
Halenbeck RF, Huang MM, Kochkocsa S, Halshan L, Linemann T;  
Pierce K, Wang Y, Wong JGF, Wu G, Zhang H;

WP1; 2004-348438/32.

New nucleic acid molecule for diagnosing, preventing or treating diseases  
such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
genetic, bacterial and viral diseases.

Claim 1; SEQ ID NO 3225; 428bp; English.

The present invention relates to an isolated nucleic acid molecule  
encoding a polypeptide which is believed to be cytostatic,  
antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
composition and methods are useful for diagnosing, preventing and  
treating diseases such as proliferative (e.g. cancer), inflammatory,  
immune, metabolic, genetic, bacterial and viral diseases. The present  
sequence represents a human secreted protein. The present sequence is  
available on WIPOMB and is not in the specification.



```

Query Match          69.2%; Score 18; DB 8; Length 525;
Best Local Similarity 20.0%; Pred. NO. 4.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      1 CXXXXXXXXXC 10
        |             |
        475 CAAATTTAC 484

Db

RESULT 37
ADP30855
ID      ADP30855 standard; protein; 549 AA.
AC      ADP30855;
XX
XX      12-AUG-2004 (first entry)
DT
XX
DE      Human secreted protein SEQ ID #1622.
XX
KM      Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KM      cancer; inflammatory; immune; human secreted protein.
OS      Homo sapiens.
XX
XX      WO2004035732-A2.
XX
XX      29-APR-2004.
PD
XX
XX      28-AUG-2003; 2003WO-US026780.
XX
XX      29-AUG-2002; 2002US-0406576P.
XX      29-AUG-2002; 2002US-0406579P.
XX      29-AUG-2002; 2002US-0406585P.
XX      29-AUG-2002; 2002US-0406588P.
XX      29-AUG-2002; 2002US-0406608P.
XX      29-AUG-2002; 2002US-0406611P.
XX      29-AUG-2002; 2002US-0406612P.
XX      29-AUG-2002; 2002US-0406616P.
XX      29-AUG-2002; 2002US-0406640P.
XX      29-AUG-2002; 2002US-0406642P.
XX      29-AUG-2002; 2002US-0406646P.
XX      29-AUG-2002; 2002US-0406653P.
XX      29-AUG-2002; 2002US-0406655P.
XX      29-AUG-2002; 2002US-0406666P.
XX      17-SEP-2002; 2002US-0410946P.
XX      17-SEP-2002; 2002US-0410947P.
XX      17-SEP-2002; 2002US-0410948P.
XX      17-SEP-2002; 2002US-0410949P.
XX      17-SEP-2002; 2002US-0410951P.
XX      17-SEP-2002; 2002US-0410953P.
XX      17-SEP-2002; 2002US-0410957P.
XX      17-SEP-2002; 2002US-0410958P.
XX      17-SEP-2002; 2002US-0410959P.
XX      17-SEP-2002; 2002US-0410960P.
XX      17-SEP-2002; 2002US-0410961P.
XX      17-SEP-2002; 2002US-0410962P.
XX      17-SEP-2002; 2002US-0411019P.
XX      17-SEP-2002; 2002US-0411022P.
XX      17-SEP-2002; 2002US-0411023P.
XX      17-SEP-2002; 2002US-0411024P.
XX      17-SEP-2002; 2002US-0411025P.
XX      17-SEP-2002; 2002US-0411035P.
XX      17-SEP-2002; 2002US-0411037P.
XX      17-SEP-2002; 2002US-0411041P.
XX      17-SEP-2002; 2002US-0411045P.
XX      17-SEP-2002; 2002US-0411046P.
XX      17-SEP-2002; 2002US-0411048P.
XX      17-SEP-2002; 2002US-0411052P.
XX      17-SEP-2002; 2002US-0411055P.
XX      17-SEP-2002; 2002US-0411073P.
XX      17-SEP-2002; 2002US-0411082P.
XX      17-SEP-2002; 2002US-0411101P.
XX      17-SEP-2002; 2002US-0411111P.

```

```

PR      18-APR-2003; 2003US-0463700P.
PR      18-APR-2003; 2003US-0463708P.
PR      18-APR-2003; 2003US-0463716P.
PR      18-APR-2003; 2003US-0463732P.
PR      02-MAY-2003; 2003US-0467199P.
PR      02-MAY-2003; 2003US-0467201P.
PR      02-MAY-2003; 2003US-0467203P.
PR      02-MAY-2003; 2003US-0467203P.
PR      19-MAY-2003; 2003US-0471306P.
PR      19-MAY-2003; 2003US-0471336P.
PR      22-MAY-2003; 2003US-0472420P.
PR      22-MAY-2003; 2003US-0472430P.
PR      09-JUN-2003; 2003US-0476609P.
PR      09-JUN-2003; 2003US-0476641P.
PR      08-JUL-2003; 2003US-0485218P.
PR      08-JUL-2003; 2003US-0485223P.
PR      08-JUL-2003; 2003US-0485224P.
PR      08-JUL-2003; 2003US-0485325P.
PR      14-JUL-2003; 2003US-0486446P.
PR      14-JUL-2003; 2003US-0486480P.
PR      15-JUL-2003; 2003US-0486891P.
PR      15-JUL-2003; 2003US-0486960P.
PR      08-AUG-2003; 2003US-0493341P.
PR      08-AUG-2003; 2003US-0493370P.
PR      08-AUG-2003; 2003US-0493573P.
PR      08-AUG-2003; 2003US-0493577P.
XX
XX      (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX      Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
XX      Halenbeck RF, Huang MM, Kothakota S, Halsehan L, Himmemann T;
XX      Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX      WPI; 2004-346438/32.
XX
XX      New nucleic acid molecule for diagnosing, preventing or treating diseases
XX      such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX      genetic, bacterial and viral diseases.
XX
XX      Claim 1; SEQ ID NO 2853; 428pp; English.
XX
XX      The present invention relates to an isolated nucleic acid molecule
XX      encoding a polypeptide which is believed to be cytostatic,
XX      antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX      composition and methods are useful for diagnosing, preventing and
XX      treating diseases such as proliferative (e.g. cancer), inflammatory,
XX      immune, metabolic, genetic, bacterial and viral diseases. The present
XX      sequence represents a human secreted protein. The present sequence is
XX      available on WIPOMEB and is not in the specification.
XX
XX      Sequence 549 AA;

Query Match          69.2%; Score 18; DB 8; Length 549;
Best Local Similarity 20.0%; Pred. NO. 4.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      1 CXXXXXXXXXC 10
        |             |
        220 CAAATTTATC 229

Db

RESULT 38
ADP31416
ID      ADP31416 standard; protein; 555 AA.
AC      ADP31416;
XX
XX      12-AUG-2004 (first entry)
DT
XX
DE      Human secreted protein SEQ ID #2183.
XX
XX      Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX      cancer; inflammatory; immune; human secreted protein.

```

[illegible]

PR	08-AUG-2003;	2003US-0493341P.	
PR	08-AUG-2003;	2003US-0493370P.	
PR	08-AUG-2003;	2003US-0493573P.	
PR	08-AUG-2003;	2003US-0493577P.	
XX			
XX			
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.		
XX			
PI	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,		
PI	Halsbeck RF, Huang MM, Kotnakota S, Haisan L, Linnemann T;		
PI	Pierce K, Wang Y, Wong UGP, Wu G, Zhang H;		
DR	WPI, 2004-348438/32.		
XX			
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases		
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,		
PT	genetic, bacterial and viral diseases.		
PS	Claim 1; SEQ ID NO 3414; 428bp; English.		
XX			
CC	The present invention relates to an isolated nucleic acid molecule		
CC	encoding a polypeptide which is believed to be cytostatic,		
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The		
CC	composition and methods are useful for diagnosing, preventing and		
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,		
CC	immune, metabolic, genetic, bacterial and viral diseases. The present		
CC	sequence represents a human secreted protein. The present sequence is		
CC	available on WIPOMB and is not in the specification.		
XX			
SQ	Sequence 555 AA;		
	Query Match	69.2%;	Score 18; DB 8; Length 555;
	Best Local Similarity	20.0%;	Pred. No. 4.6;
	Matches 2; Conservative	0;	Mismatches 8; Indels 0; Gaps 0
Qy	1 CXXXXXXXXXC 10		
Db	211 CTTATTTC 220		
RESULT 39			
ID	ADP31417		
ID	ADP31417 standard; protein; 555 AA.		
XX			
AC	ADP31417;		
XX			
DT	12-AUG-2004 (first entry)		
XX			
DE	Human secreted protein SEQ ID #2184.		
XX			
KM	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;		
XX	Cancer; Inflammatory; Immune; human secreted protein.		
OS	Homo sapiens.		
XX			
PN	MO2004035732-A2.		
PD			
XX	29-APR-2004.		
PF			
XX	28-AUG-2003; 2003WO-US026780.		
XX			
PR	29-AUG-2002; 2002US-0406576P.		
PR	29-AUG-2002; 2002US-0406579P.		
PR	29-AUG-2002; 2002US-0406585P.		
PR	29-AUG-2002; 2002US-0406588P.		
PR	29-AUG-2002; 2002US-0406608P.		
PR	29-AUG-2002; 2002US-0406611P.		
PR	29-AUG-2002; 2002US-0406612P.		
PR	29-AUG-2002; 2002US-0406616P.		
PR	29-AUG-2002; 2002US-0406640P.		
PR	29-AUG-2002; 2002US-0406642P.		
PR	29-AUG-2002; 2002US-0406646P.		
PR	29-AUG-2002; 2002US-0406653P.		
PR	29-AUG-2002; 2002US-0406655P.		

PR	29-AUG-2002;	2002US-0406666P.
PR	17-SEP-2002;	2002US-0410946P.
PR	17-SEP-2002;	2002US-0410947P.
PR	17-SEP-2002;	2002US-0410948P.
PR	17-SEP-2002;	2002US-0410949P.
PR	17-SEP-2002;	2002US-0410953P.
PR	17-SEP-2002;	2002US-0410957P.
PR	17-SEP-2002;	2002US-0410958P.
PR	17-SEP-2002;	2002US-0410959P.
PR	17-SEP-2002;	2002US-0410960P.
PR	17-SEP-2002;	2002US-0410961P.
PR	17-SEP-2002;	2002US-0410962P.
PR	17-SEP-2002;	2002US-0411019P.
PR	17-SEP-2002;	2002US-0411022P.
PR	17-SEP-2002;	2002US-0411023P.
PR	17-SEP-2002;	2002US-0411024P.
PR	17-SEP-2002;	2002US-0411032P.
PR	17-SEP-2002;	2002US-0411035P.
PR	17-SEP-2002;	2002US-0411037P.
PR	17-SEP-2002;	2002US-0411041P.
PR	17-SEP-2002;	2002US-0411045P.
PR	17-SEP-2002;	2002US-0411046P.
PR	17-SEP-2002;	2002US-0411048P.
PR	17-SEP-2002;	2002US-0411101P.
PR	17-SEP-2002;	2002US-0411111P.
PR	18-APR-2003;	2003US-0463700P.
PR	18-APR-2003;	2003US-0463708P.
PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US-0463732P.
PR	02-MAY-2003;	2003US-0467239P.
PR	02-MAY-2003;	2003US-0467201P.
PR	02-MAY-2003;	2003US-0467203P.
PR	02-MAY-2003;	2003US-0467230P.
PR	19-MAY-2003;	2003US-0471306P.
PR	19-MAY-2003;	2003US-0471366P.
PR	22-MAY-2003;	2003US-0472420P.
PR	22-MAY-2003;	2003US-0472430P.
PR	09-JUN-2003;	2003US-0476609P.
PR	09-JUN-2003;	2003US-0476641P.
PR	08-JUL-2003;	2003US-0485218P.
PR	08-JUL-2003;	2003US-0485223P.
PR	08-JUL-2003;	2003US-0485224P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486446P.
PR	14-JUL-2003;	2003US-0486448P.
PR	15-JUL-2003;	2003US-0486891P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX		
XX	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,	
PI	Halenebeck RF, Huang MM, Kochakota S, Hsiehan L, Linemann T,	
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;	
XX		
XX	WPI; 2004-348438/32.	
DR		
XX		
XX	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	genetic, bacterial and viral diseases.	
XX		
XX	Claim 1; SEQ ID NO 3415; 428bp; English.	
PS		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	

CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPoWEB and is not in the specification.  
XX  
SQ Sequence 555 AA;

Query Match	69.2%	Score 18;	DB 8;	Length 55;
Best Local Similarity	20.0%	Pred. No. 4.6;		
Matches	2;	Conservative	0;	Mismatches 8;
			Indels	0;
			Gaps	0

Qy 1 CXXXXXXXXC 10  
Db 211 CTTATTATAC 220

RESULT 40  
ADP31699  
ID ADP31699 standard; protein; 588 AA

AC	ADP31699;
XX	
DT	12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2466.

KM Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide,  
KM cancer; inflammatory; immune; human secreted protein.  
KM

OS Homo sapiens.

PN WO2004035732-A2

PD 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

PR 29-AUG-2002; 2002US-040657

PR 29-AUG-2002; 2002US-040658

PR 29-AUG-2002; 2002US-040660

PR 29-AUG-2002; 2002US-040661

PR 29-AUG-2002; 2002US-040664

PR 29-AUG-2002; 2002US-040664

PR 29-AUG-2002; 2002US-040665

PR 17-SEP-2002; 2002US-041094

PR 17-SEP-2002; 2002US-041094

PR 17-SEP-2002; 2002US-041095

PR 17-SEP-2002; 2002US-041095

PR 17-SEP-2002; 2002US-041096

PR 17-SEP-2002; 2002US-041096

PR 17-SEP-2002; 2002US-041102

PR 17-SEP-2002; 2002US-041102

PR 17-SEP-2002; 2002US-041103

PR 17-SEP-2002; 2002US-041104

PR 17-SEP-2002; 2002US-041104

PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476039P.  
PR 09-JUN-2003; 2003US-0476041P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Hsiehan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.

PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT genetic, proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX Claim 1; SEQ ID NO 3697; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.

XX Sequence 588 AA;

Query Match 69.2%; Score 18; DB 8; Length 588;  
Best Local Similarity 20.0%; Pred. No. 4.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10  
Db 277 CATTTTTC 286

RESULT 41  
ADZ56209 standard; protein; 591 AA.  
XX ADZ56209;  
AC ADZ56209;

XX 30-JUN-2005 (first entry)  
DT Human KIAA0779 splice variant clone 24980850:24980849a, nonTM\_1 #1.  
XX KIAA0779; gene expression; inflammation; antiinflammatory; cancer;  
XX proliferation; neoplasia; cytostatic; immune disorder; immunomodulator;  
XX metabolic disorder; metabolic; viral infection; virucide; infection.  
OS Homo sapiens.  
PN WO2005035569-A2.  
PD 21-APR-2005.  
PF 12-OCT-2004; 2004WO-US033408.  
XX 10-OCT-2003; 2003US-0510612P.  
PR (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PA Wong JGP, Hestir K, Collins AL;  
PI Wong JGP, Hestir K, Collins AL;  
XX WPI; 2005-296268/30.  
DR New isolated KIAA0779 nucleic acids and polypeptides, useful for  
PT diagnosing, preventing and/or treating inflammatory, immune, viral  
PT disorders and cancer, such as kidney, lung, ovarian, bladder, breast,  
PT prostate and skin cancers.  
XX Claim 1; SEQ ID NO 14; 121bp; English.  
XX The invention relates to an isolated nucleic acid molecule comprises at  
CC least one polynucleotide sequence (appearing as ADZ56196-ADZ56197,  
CC ADZ56199-ADZ56200, ADZ56226-ADZ56227 and ADZ56230-ADZ56231), sequences  
CC hybridizing to them under high stringency conditions, sequences having at  
CC least 80% sequence identity to them, their complements or biologically  
CC active fragments. The nucleic acids are splice variants of the human  
CC KIAA0779 gene, encoding transmembrane domain protein(s). Also included  
CC are a double-stranded isolated nucleic acid molecule comprising the  
CC nucleic acid molecule cited above, a vector comprising the isolated  
CC nucleic acid molecule cited above, (and a promoter that regulates the  
CC expression of the nucleic acid molecule), a recombinant host cell  
CC comprising the nucleic acid molecule cited above, an isolated polypeptide  
CC (comprising the encoded proteins or the non-transmembrane (TM) regions of  
CC the proteins), a method of making a recombinant host cell, a method of  
CC making a polypeptide, a method of determining the presence of the nucleic  
CC acid molecule cited above in a sample, a method of determining the  
CC presence of a specific antibody to the polypeptide of (4) in a sample, a  
CC method of determining the presence of the polypeptide in a sample, an  
CC antibody specifically binding to and/or interfering with the biological  
CC activity of the nucleic acid molecule cited above (or the polypeptide or  
CC its biologically active fragment), a composition comprising a  
CC pharmaceutical carrier or excipient (and one or more active agents chosen  
CC from the nucleic acid molecule cited above, the vector, the polypeptide,  
CC and the antibody), a bacteriophage comprising the antibody (or its  
CC fragment), a bacterial cell comprising the bacteriophage, a recombinant  
CC host cell that produces the antibody, an animal injected with one or more  
CC active agents (chosen from the nucleic acid molecule, the vector, the  
CC host cell, the polypeptide, and the antibody), a diagnostic kit  
CC (comprising a nucleic acid molecule having at least 6 contiguous  
CC nucleotides from the nucleic acid molecule cited above, the isolated  
CC polypeptide, or antibody, and reagents to carry out an immunoassay), a  
CC method of making an antibody, a method of identifying a modulating agent  
CC that modulates the biological activity of the polypeptide, a modulator  
CC composition comprising a modulator and a pharmaceutical carrier, a method  
CC of treating a disease in a subject, and a method of treating cancer,  
CC preferably kidney, cervical, squamous lung, ovarian, bladder, breast,  
CC endometrial, prostate, and skin cancer, in a subject. The methods and  
CC compositions of the present invention are useful for the diagnosis,  
CC prevention and/or treatment of inflammatory, immune, viral disorders and  
CC cancer, such as kidney, cervical, squamous lung, ovarian, bladder,  
CC breast, endometrial, prostate, and skin cancer. The present sequence

CC represents a non TM region of a protein encoded by a cDNA splice variant  
 CC from the KIAA0779 gene.  
 XX  
 SQ Sequence 591 AA;

Query Match 69.2%; Score 18; DB 9; Length 591;  
 Best Local Similarity 20.0%; Pred. No. 4.6;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 10  
 DB 173 CAAAAAAAC 182

RESULT 42  
 AD256223  
 ID AD256223 standard; protein; 591 AA.

AC AD256223;  
 DT 30-JUN-2005 (first entry)

DE Human KIAA0779 splice variant clone 24980850:24980849a, nonTM\_1 #2.

KW KIAA0779; gene expression; inflammation; antiinflammatory; cancer;  
 KW proliferation; neoplasm; cytoskeletal; immune disorder; immunomodulator;  
 KW metabolic disorder; metabolic; viral infection; virucide; infection.

OS Homo sapiens.

PN WO2005035569-A2.

PD 21-APR-2005.

PE 12-OCT-2004; 2004WO-US033408.

PR 10-OCT-2003; 2003US-0510612P.

PA (FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Wong JGP, Heetir K, Collins ALT;

DR WPI; 2005-296268/30.

PT New isolated KIAA0779 nucleic acids and polypeptides, useful for  
 PT diagnosing, preventing and/or treating inflammatory, immune, viral  
 PT disorders and cancer, such as kidney, lung, ovarian, bladder, breast,  
 PT prostate and skin cancers.

PS Claim 11; SEQ ID NO 28; 121bp; English.

CC The invention relates to an isolated nucleic acid molecule comprises at  
 CC least one polynucleotide sequence (appearing as AD256196-AD256197,  
 CC AD256199-AD256200, AD256226-AD256227 and AD256230-AD256231), sequences  
 CC hybridizing to them under high stringency conditions, sequences having at  
 CC least 80% sequence identity to them, their complements or biologically  
 CC active fragments. The nucleic acids are splice variants of the human  
 CC KIAA0779 gene, encoding transmembrane domain protein(s). Also included  
 CC are a double-stranded isolated nucleic acid molecule comprising the  
 CC nucleic acid molecule cited above, a vector comprising the isolated  
 CC nucleic acid molecule cited above, (and a promoter that regulates the  
 CC expression of the nucleic acid molecule), a recombinant host cell  
 CC comprising the nucleic acid molecule cited above, an isolated polypeptide  
 CC (comprising the encoded proteins or the non-transmembrane (TM) regions of  
 CC the proteins), a method of making a recombinant host cell, a method of  
 CC making a polypeptide, a method of determining the presence of the nucleic  
 CC acid molecule cited above in a sample, a method of determining the  
 CC presence of a specific antibody to the polypeptide of (4) in a sample, a  
 CC method of determining the presence of the polypeptide in a sample, an  
 CC antibody specifically binding to and/or interfering with the biological  
 CC activity of the nucleic acid molecule cited above (or the polypeptide or  
 CC its biologically active fragment), a composition comprising a  
 CC pharmaceutical carrier or excipient (and one or more active agents chosen

CC from the nucleic acid molecule cited above, the vector, the polypeptide,  
 CC and the antibody), a bacteriophage comprising the antibody (or its  
 CC fragment), a bacterial cell comprising the bacteriophage, a recombinant  
 CC host cell that produces the antibody, an animal injected with one or more  
 CC active agents (chosen from the nucleic acid molecule, the vector, the  
 CC host cell, the polypeptide, and the antibody), a diagnostic kit  
 CC (comprising a nucleic acid molecule having at least 6 contiguous  
 CC nucleotides from the nucleic acid molecule cited above, the isolated  
 CC polypeptide, or antibody, and reagents to carry out an immunoassay), a  
 CC method of making an antibody, a method of identifying a modulating agent  
 CC that modulates the biological activity of the polypeptide, a modulator  
 CC composition comprising a modulator and a pharmaceutical carrier, a method  
 CC of treating a disease in a subject, and a method of treating cancer,  
 CC preferably kidney, cervical, squamous lung, ovarian, bladder, breast,  
 CC endometrial, prostate, and skin cancer, in a subject. The methods and  
 CC compositions of the present invention are useful for the diagnosis,  
 CC prevention and/or treatment of inflammatory, immune, viral disorders and  
 CC cancer, such as kidney, cervical, squamous lung, ovarian, bladder,  
 CC breast, endometrial, prostate, and skin cancer. The present sequence  
 CC represents a non TM region of a protein encoded by a cDNA splice variant  
 CC from the KIAA0779 gene.

SQ Sequence 591 AA;

Query Match 69.2%; Score 18; DB 9; Length 591;  
 Best Local Similarity 20.0%; Pred. No. 4.6;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXC 10  
 DB 173 CAAAAAAAC 182

RESULT 43

ADP30865  
 ID ADP30865 standard; protein; 600 AA.

AC ADP30865;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1632.

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
 KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN WO2004035732-A2.

PD 29-APR-2004.

PE 28-AUG-2003; 2003WO-US026780.

PR 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.

XX	PR	17-SEP-2002;	2002US-0410957P.
XX	PR	17-SEP-2002;	2002US-0410958P.
XX	PR	17-SEP-2002;	2002US-0410959P.
XX	PR	17-SEP-2002;	2002US-0410960P.
XX	PR	17-SEP-2002;	2002US-0410961P.
XX	PR	17-SEP-2002;	2002US-0410962P.
XX	PR	17-SEP-2002;	2002US-0411019P.
XX	PR	17-SEP-2002;	2002US-0411022P.
XX	PR	17-SEP-2002;	2002US-0411023P.
XX	PR	17-SEP-2002;	2002US-0411024P.
XX	PR	17-SEP-2002;	2002US-0411032P.
XX	PR	17-SEP-2002;	2002US-0411035P.
XX	PR	17-SEP-2002;	2002US-0411037P.
XX	PR	17-SEP-2002;	2002US-0411041P.
XX	PR	17-SEP-2002;	2002US-0411045P.
XX	PR	17-SEP-2002;	2002US-0411046P.
XX	PR	17-SEP-2002;	2002US-0411048P.
XX	PR	17-SEP-2002;	2002US-0411052P.
XX	PR	17-SEP-2002;	2002US-0411055P.
XX	PR	17-SEP-2002;	2002US-0411073P.
XX	PR	17-SEP-2002;	2002US-0411082P.
XX	PR	17-SEP-2002;	2002US-0411101P.
XX	PR	17-SEP-2002;	2002US-0411111P.
XX	PR	18-APR-2003;	2003US-0463700P.
XX	PR	18-APR-2003;	2003US-0463708P.
XX	PR	18-APR-2003;	2003US-0463716P.
XX	PR	18-APR-2003;	2003US-0463732P.
XX	PR	02-MAY-2003;	2003US-0467199P.
XX	PR	02-MAY-2003;	2003US-0467201P.
XX	PR	02-MAY-2003;	2003US-0467203P.
XX	PR	02-MAY-2003;	2003US-0467230P.
XX	PR	19-MAY-2003;	2003US-0471306P.
XX	PR	19-MAY-2003;	2003US-0471336P.
XX	PR	22-MAY-2003;	2003US-0472420P.
XX	PR	22-MAY-2003;	2003US-0472430P.
XX	PR	09-JUN-2003;	2003US-0476609P.
XX	PR	09-JUN-2003;	2003US-0476641P.
XX	PR	08-JUL-2003;	2003US-0485218P.
XX	PR	08-JUL-2003;	2003US-0485223P.
XX	PR	08-JUL-2003;	2003US-0485224P.
XX	PR	08-JUL-2003;	2003US-0485325P.
XX	PR	14-JUL-2003;	2003US-0486446P.
XX	PR	14-JUL-2003;	2003US-0486480P.
XX	PR	15-JUL-2003;	2003US-0486919P.
XX	PR	15-JUL-2003;	2003US-0486960P.
XX	PR	08-AUG-2003;	2003US-0493341P.
XX	PR	08-AUG-2003;	2003US-0493370P.
XX	PR	08-AUG-2003;	2003US-0493573P.
XX	PR	08-AUG-2003;	2003US-0493577P.
XX	PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX	PI	Williams LT,	Chiu K, Lee E, Hestir K, Beurang PA, Behrens D,
XX	PI	Hallenbeck RF,	Huang MM, Kochakota S, Haishan L, Linemann T;
XX	PI	Pierce K,	Wang Y, Wong JGP, Wu G, Zhang H;
XX	DR	WPI,	2004-348438/32.
PT	XX	New nucleic acid molecule for diagnosing, preventing or treating diseases	
PT	XX	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
PT	XX	genetic, bacterial and viral diseases.	
PS	XX	Claim 1; SEQ ID NO 2863; 428bp; English.	
CC	XX	The present invention relates to an isolated nucleic acid molecule	
CC	XX	encoding a polypeptide which is believed to be cytostatic,	
CC	XX	anti-inflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	XX	composition and methods are useful for diagnosing, preventing and	
CC	XX	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	XX	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	XX	sequence represents a human secreted protein. The present sequence is	
CC	XX	available on WIPOMB and is not in the specification.	

[illegible]

PR 17-SEP-2002; 2002US-041111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485255P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.

## (FIVE-) FIVE PRIME THERAPEUTICS INC.

PI Williams LR, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haisman L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PI genetic, bacterial and viral diseases.

PS Claim 1; SEQ ID NO 2938; 428pp; English.

XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
XX available on WIPOWEB and is not in the specification.

SQ Sequence 604 AA;

Query Match 69.2%; Score 18; DB 8; Length 604;

Best Local Similarity 20.0%; Pred. No. 4.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10  
DB 430 CTTATTAAAC 439

## RESULT 45

ADP30941  
ID ADP30941 standard; protein; 604 AA.

XX ADP30941;

XX 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1708.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

XX WO2004035732-A2.

XX 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

PR 29-AUG-2002; 2002US-0405576P.

PR 29-AUG-2002; 2002US-0405579P.

PR 29-AUG-2002; 2002US-0405585P.

PR 29-AUG-2002; 2002US-0405588P.

PR 29-AUG-2002; 2002US-0406088P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406640P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410961P.

PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.

PR 17-SEP-2002; 2002US-0411035P.

PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.

PR 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411101P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467199P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 22-MAY-2003; 2003US-0472430P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 08-JUL-2003; 2003US-0485325P.

PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2939; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 604 AA;

Query Match 69.2%; Score 18; DB 8; Length 604;  
Best Local Similarity 20.0%; Pred. No. 4.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10  
DB 430 CTTATTAAAC 439

RESULT 46  
ADP30507  
ID ADP30507 standard; protein; 605 AA.  
XX  
AC ADP30507;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1274.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406662P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2505; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.



CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.

SO Sequence 605 AA;

Query Match 69.2%; Score 18; DB 8; Length 605;  
Best Local Similarity 20.0%; Pred. No. 4.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10  
DB 570 CAATTATTC 579

RESULT 47  
ADP31657

ID ADP31657 standard; protein; 617 AA.

AC ADP31657;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #2424.

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN WO2004035732-A2.

XX WO2004035732-A2.

XX 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

PR 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410957P.

PR 17-SEP-2002; 2002US-0411046P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411082P.

PR 17-SEP-2002; 2002US-0411101P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.

PR 18-APR-2003; 2003US-0463732P.

PR 02-MAY-2003; 2003US-0467119P.

PR 02-MAY-2003; 2003US-0467201P.

PR 02-MAY-2003; 2003US-0467203P.

PR 02-MAY-2003; 2003US-0467230P.

PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.

PR 09-JUN-2003; 2003US-0476609P.

PR 09-JUN-2003; 2003US-0476641P.

PR 08-JUL-2003; 2003US-0485218P.

PR 08-JUL-2003; 2003US-0485223P.

PR 08-JUL-2003; 2003US-0485224P.

PR 14-JUL-2003; 2003US-0486446P.

PR 14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.

PR 15-JUL-2003; 2003US-0486960P.

PR 08-AUG-2003; 2003US-0493341P.

PR 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.

PR 08-AUG-2003; 2003US-0493577P.

XX (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 3655; 428pp; English.

XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.

XX Sequence 617 AA;

Query Match 69.2%; Score 18; DB 8; Length 617;  
Best Local Similarity 20.0%; Pred. No. 4.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10  
DB 67 CAATTATTC 76

RESULT 48

ID ABR41596 standard; protein; 626 AA.

XX ABR41596

AC ABR1596;  
XX  
XX 02-JUN-2003 (first entry)  
DT  
XX Human DITHP antigen recognition protein.  
DE  
XX  
XX Human, dithp, diagnostic and therapeutic polynucleotide; diagnosis;  
KM cancer; cell proliferative disorder; autoimmune disorder;  
KM inflammatory disorder; infection; hormonal disorder; metabolic disorder;  
KM neurological disorder; gastrointestinal disorder; transport disorder;  
KM connective tissue disorder; drug screening; proteome analysis;  
KM gene therapy; antisense therapy; genotyping; transgenic animal; knock in;  
KM disease model; toxicological testing; transcript imaging;  
KM antigen recognition.  
XX  
XX Homo sapiens.  
OS  
XX WO200297031-A2.  
XX  
XX 05-DEC-2002.  
XX  
XX 27-MAR-2002; 2002WO-US010056.  
XX  
XX 28-MAR-2001; 2001US-0279619P.  
PR 29-MAR-2001; 2001US-0280067P.  
PR 29-MAR-2001; 2001US-0280068P.  
PR 16-MAY-2001; 2001US-0291280P.  
PR 17-MAY-2001; 2001US-0291829P.  
PR 17-MAY-2001; 2001US-0291849P.  
PR 19-JUN-2001; 2001US-0299428P.  
PR 20-JUN-2001; 2001US-0299776P.  
PR 20-JUN-2001; 2001US-0300001P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
PI Dufour GE, Hillman JL, Yu JY, Thason O, Yap PE, Amesby SR;  
PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kiesefeld Y, Geertsen ER;  
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
PI Flores V, Marwaha R, Lo A, Ian RV, Urashka ME;  
XX  
XX WPI; 2003-129518/12.  
XX  
XX N-PSDB; ACC46534.  
XX  
XX Novel human diagnostic and therapeutic polypeptide useful for identifying  
PT test compound which specifically binds to a polypeptide encoded by human  
DT diagnostic and therapeutic polynucleotide, and to induce antibodies.  
XX  
XX Claim 27; SEQ ID NO 1131, 591pp; English.  
XX  
XX The invention relates to novel human diagnostic and therapeutic  
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded  
CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to  
CC polynucleotide sequences at least 90% identical to the dithp cDNA  
CC sequences of the invention; recombinant vectors, host cells and  
CC transgenic organisms comprising a dithp nucleic acid sequence; the  
CC recombinant production of DITHP proteins; antibodies specific for DITHP  
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of  
CC detecting dithp nucleotide and protein sequences; methods of screening  
CC for compounds which specifically bind a DITHP protein; and methods of  
CC assessing the toxicity of test compounds using a dithp hybridisation  
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the  
CC diagnosis of a wide variety of conditions including cancer and other cell  
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,  
CC viral, fungal or parasitic infections; hormonal disorders; metabolic  
CC disorders; neurological disorders; gastrointestinal disorders; transport  
CC disorders; and connective tissue disorders. They may also be used to  
CC screen for modulators of protein activity or gene expression. DITHP  
CC proteins can additionally be used in analysis of the proteome of a tissue  
CC or cell type and to induce antibodies. The dithp nucleic acids are  
CC additionally useful in somatic or germline gene therapy of the disorders  
CC mentioned above, as a source of antisense sequences, as a source of  
CC probes and primers, in genotyping and identification of individuals, in

CC the generation of transgenic animal models of human disease or knock in  
CC humanised animals, in toxicological testing, and in transcript imaging.  
CC The present sequence represents a DITHP protein which has antigen  
CC recognition activity. Note: The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 626 AA;  
XX  
XX Query Match 69.2%; Score 18; DB 6; Length 626;  
XX Best Local Similarity 20.0%; Pred. No. 4.7;  
XX Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
XX  
QY 1 CXXXXXXXC 10  
XX  
DB 41 CTSTSSAASC 50  
XX  
RESULT 49  
ADP30858  
ID ADP30858 standard; protein; 645 AA.  
XX  
XX ADP30858;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #1625.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KM cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
XX  
XX WO200403732-A2.  
XX  
XX 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411012P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411101P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463700P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471366P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.

## (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;  
 PI Halebek R, Huang MM, Kochakota S, Haislan L, Linnemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 2856; 428pp; English.

XX The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMB and is not in the specification.

XX Sequence 645 AA;

XX Query Match 69.2%; Score 18; DB 8; Length 645;

XX Best Local Similarity 20.0%; Pred. No. 4.7;  
 XX Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10  
 DB 161 CAAATTTTTC 170

RESULT 50  
 ID AAY94907 standard; protein; 653 AA.  
 XX

AC AAY94907;  
 DT 16-JUN-2000 (first entry)

DE Human secreted protein clone cal06\_19x protein sequence SEQ ID NO:20.

XX Human; secreted protein; immunostimulant; immunosuppressant; virucide;  
 KW antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;  
 KW antidiabetic; antiaesthetic; antirheumatic; antirheumatoid; protozoacide;  
 KW antithyroid; immune deficiency; severe combined immunodeficiency; SCID;  
 KW infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;  
 KW connective tissue disease; multiple sclerosis; erythematosis;  
 KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;  
 KW Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;  
 KW insulin dependent diabetes mellitus; graft-versus-host-disease;  
 KW autoimmune inflammatory eye disease; allergy.

XX Homo sapiens.

OS WO200009552-A1.

PN 24-FEB-2000.

PD 13-AUG-1999; 99WO-US018298.

XX 14-AUG-1998; 98US-0096622P.  
 XX 17-AUG-1998; 98US-0096815P.  
 XX 04-SEP-1998; 98US-0099229P.  
 XX 23-OCT-1998; 98US-0105368P.  
 XX 08-JAN-1999; 99US-011534P.  
 XX 12-FEB-1999; 99US-0119931P.  
 XX 18-FEB-1999; 99US-0120575P.  
 XX 30-APR-1999; 99US-0132020P.  
 XX 11-AUG-1999; 99US-0148424P.

## (GEMT) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;  
 PI Mong CG, Clark HF, Fechtel K;  
 DR WPI; 2000-205979/18.

XX New polynucleotides encoding secreted proteins, which may have e.g.  
 PT nutritional, chemokine, immune stimulating or suppressing, hematopoiesis  
 PT regulating, tissue growth, activin/inhibin antiinflammatory or tumor  
 PT inhibition activity.

XX Claim 29; Page 487-489; 641pp; English.

XX AAA1618 to AAA1697 encode the human secreted proteins given in AAY94898  
 CC to AAY94980, isolated from human adult brain, adult thymoid, adult  
 CC retina, foetal carcinoma, adult blood, adult neural, foetal kidney, adult  
 CC placenta, adult testis, whole embryo, adult cartilage, kidney, foetal  
 CC brain, adult thymus, foetal placenta, adult uterus, adult tumour, and  
 CC adult bladder, cDNA libraries. The polynucleotides and proteins are  
 CC predicted to have biological activities which would make them suitable  
 CC for treating, preventing or ameliorating medical conditions in humans and  
 CC animals. The polynucleotides can be used as markers for tissues in which  
 CC the protein is preferentially expressed, as molecular weight markers on  
 CC Southern gels, and as chromosome markers or tags to identify chromosomes  
 CC or to map gene positions. The proteins can be used in the treatment of  
 CC immune deficiencies and disorders, such as severe combined  
 CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other  
 CC infections. These infections include human immunodeficiency virus (HIV),  
 CC hepatitis, herpesviruses, mycobacteria, leishmania spp., malaria and  
 CC candidiasis. The proteins can be used to treat autoimmune disorders such  
 CC as connective tissue disease, multiple sclerosis, systemic lupus  
 CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,  
 CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent  
 CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and  
 CC autoimmune inflammatory eye disease. The proteins can also be used to  
 CC treat allergic conditions, such as asthma. AAA1698 to AAA16774 represent

CC probes for the human secreted proteins from the present invention  
XX  
SQ Sequence 653 AA;

Query Match 69.2%; Score 18; DB 3; Length 653;  
Best Local Similarity 20.0%; Pred. No. 4.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXC 10  
|  
Db 173 CAAAAAAC 182

Search completed: January 4, 2006, 15:56:17  
Job time : 70.4348 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 10.3478 Seconds  
(without alignments)  
92.983 Million cell updates/sec

Title: US-09-932-322-11  
Perfect score: 26  
Sequence: 1 CXXXXXXXXC 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 1000 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	18	69.2	313	2	S59448	hypothetical prote
2	18	69.2	480	2	E75433	hypothetical prote
3	17	65.4	194	2	JQ1560	hypothetical 20.6K
4	17	65.4	220	2	JC4082	coat protein - Cym
5	17	65.4	376	2	AB2920	cobalamin biosynth
6	17	65.4	388	2	C97694	cobalamin biosynth
7	17	65.4	389	2	T23167	hypothetical prote
8	17	65.4	570	2	T37314	probable kexin (EC
9	17	65.4	634	2	S33575	dnak-type molecu
10	17	65.4	942	2	D87803	protein bli-4D [lm
11	17	65.4	2704	2	S09118	G surface protein
12	16	61.5	71	2	F84312	hypothetical prote
13	16	61.5	85	2	E70531	hypothetical prote
14	16	61.5	104	2	S51479	drought-induced pr
15	16	61.5	105	2	A71249	hypothetical prote
16	16	61.5	108	2	F72549	hypothetical prote
17	16	61.5	126	2	A23473	chymotrypsin-like
18	16	61.5	129	2	T49498	hypothetical prote
19	16	61.5	133	2	A96746	hypothetical prote
20	16	61.5	147	2	S09762	hypothetical prote
21	16	61.5	153	2	PN0103	hypothetical 17K p
22	16	61.5	161	2	T28088	hypothetical prote
23	16	61.5	164	2	JQ1252	hypothetical 16.7K
24	16	61.5	229	2	T34277	hypothetical prote
25	16	61.5	242	2	A45724	pectate lyase (EC
26	16	61.5	244	2	S72219	chymotrypsin B - A
27	16	61.5	251	2	A55035	cysteine-rich prote
28	16	61.5	262	1	JQ1724	BI membrane glycop
29	16	61.5	263	2	S47537	chymotrypsin (EC 3

30	16	61.5	275	2	A85856	probable elongatio
31	16	61.5	275	2	B64986	hypothetical 30.9
32	16	61.5	275	2	G91011	probable elongatio
33	16	61.5	294	2	T23682	hypothetical prote
34	16	61.5	306	2	C71498	probable yop trans
35	16	61.5	317	2	T00500	probable alicitor
36	16	61.5	321	2	A84792	hypothetical prote
37	16	61.5	349	2	JE0202	paired-box-contain
38	16	61.5	380	2	T04007	hypothetical prote
39	16	61.5	397	2	T26077	hypothetical prote
40	16	61.5	422	2	S41514	Rad52 protein homo
41	16	61.5	473	2	C81039	lipopolysaccharide
42	16	61.5	473	2	C81984	probable lipopolys
43	16	61.5	475	2	T36342	probable glutamate
44	16	61.5	486	2	B86460	hypothetical prote
45	16	61.5	510	2	A55207	glycerol-3-phospha
46	16	61.5	537	2	G31277	guanine transport
47	16	61.5	580	2	B38418	jockey protein 1 -
48	16	61.5	638	2	T22518	hypothetical prote
49	16	61.5	642	2	C81345	hypothetical prote
50	16	61.5	648	2	T37581	probable serine-th
51	16	61.5	708	1	TFBOL	lactoferrin p
52	16	61.5	708	2	JC2323	lactoferrin - goat
53	16	61.5	711	1	TFHUL	99.9K hypothetical
54	16	61.5	907	2	A86460	hypothetical prote
55	16	61.5	975	2	T29908	hypothetical prote
56	16	61.5	1031	2	T06130	hypothetical prote
57	16	61.5	1069	2	T22138	hypothetical prote
58	16	61.5	1105	2	T22132	hypothetical prote
59	16	61.5	1136	2	AB1581	different protein
60	16	61.5	1380	2	T18309	receptor-adenylate
61	16	61.5	1474	2	D88550	protein ZC84.6 [lm
62	16	61.5	2233	2	T28669	surface protein 51
63	16	61.5	2395	1	S50820	surface protein ty
64	16	61.5	2533	2	T28675	alpha-51D immobili
65	16	61.5	2533	2	T28674	alpha-51D-immobili
66	16	61.5	2543	2	T31687	surface antigen - p
67	16	61.5	2718	2	A23475	metallothionein M
68	15	57.7	40	1	SMFP	metallothionein M
69	15	57.7	40	2	B61194	metallothionein M
70	15	57.7	43	2	S33483	hypothetical prote
71	15	57.7	74	2	AB2210	hypothetical prote
72	15	57.7	93	2	T18118	hypothetical prote
73	15	57.7	109	2	S64309	hypothetical prote
74	15	57.7	111	2	A72079	hypothetical prote
75	15	57.7	111	2	E86544	hypothetical prote
76	15	57.7	113	2	T10136	hypothetical prote
77	15	57.7	123	2	B81120	probable periplasm
78	15	57.7	124	2	C81707	hypothetical prote
79	15	57.7	131	2	T36866	probable iron-sulf
80	15	57.7	135	2	T38956	very hypothetical
81	15	57.7	139	2	D97348	hypothetical prote
82	15	57.7	141	2	AE0029	hypothetical prote
83	15	57.7	141	2	PCI294	Merr-family trans
84	15	57.7	146	2	T01683	crophosolite surfac
85	15	57.7	149	2	AF2977	immobilization ant
86	15	57.7	158	2	T14975	conserved hypotet
87	15	57.7	160	2	F98305	hypothetical prote
88	15	57.7	169	1	S18946	hypothetical 16.6K
89	15	57.7	171	2	T10315	ultra high-sulfur
90	15	57.7	179	2	T22471	hypothetical prote
91	15	57.7	181	2	T31481	hypothetical prote
92	15	57.7	185	2	T51003	hypothetical prote
93	15	57.7	189	2	B90969	probable tail asse
94	15	57.7	191	2	S69735	hypothetical prote
95	15	57.7	195	2	T18617	hypothetical prote
96	15	57.7	199	2	C90734	probable outer mem
97	15	57.7	199	2	E85584	hypothetical prote
98	15	57.7	203	2	H81695	recombination prot
99	15	57.7	204	2	A28864	ShuY-like protein
100	15	57.7	207	2	A91177	hypothetical prote
101	15	57.7	207	2	B86023	hypothetical-relat
102	15	57.7	208	2	T07732	tuberculosis-relat

103	15	57.7	211	2	S34274	probable secretory
104	15	57.7	211	2	UC4883	scytallopepsin B
105	15	57.7	212	2	T22463	hypothetical prote
106	15	57.7	214	2	C70812	probable 1pqq prot
107	15	57.7	219	1	QRECTC	carbonic dehydrat
108	15	57.7	219	2	H96677	carbonic anhydrase
109	15	57.7	219	2	D85528	carbonic anhydrase
110	15	57.7	220	2	S22493	coat protein - Cym
111	15	57.7	223	2	T33194	hypothetical prote
112	15	57.7	223	2	T33193	hypothetical prote
113	15	57.7	227	2	C90899	probable tail asse
114	15	57.7	230	2	T49555	related to high cy
115	15	57.7	231	2	S46511	isoprenyl transfer
116	15	57.7	231	2	S41894	gene 3' protein -
117	15	57.7	238	1	JH0367	ribonuclease (EC 3
118	15	57.7	242	1	AB2007	transposase all160
119	15	57.7	244	2	T31838	hypothetical prote
120	15	57.7	245	1	KY80A	chymotrypsin (EC 3
121	15	57.7	245	1	KY80B	chymotrypsin (EC 3
122	15	57.7	247	2	B90997	probable tail asse
123	15	57.7	247	2	T27778	hypothetical prote
124	15	57.7	248	2	T03868	hypothetical prote
125	15	57.7	248	2	T03869	hypothetical prote
126	15	57.7	248	2	T31841	hypothetical prote
127	15	57.7	256	1	TRFF	trypsin-like prote
128	15	57.7	256	1	T27836	hypothetical prote
129	15	57.7	258	2	C86286	protein Pg1.16 [1
130	15	57.7	260	2	E90344	hydroxyethylchazo
131	15	57.7	262	1	G64977	hydroxyethylchazo
132	15	57.7	262	2	E85837	hydroxyethylchazo
133	15	57.7	262	2	C90992	hydroxyethylchazo
134	15	57.7	263	1	KYR1B	chymotrypsin (EC 3
135	15	57.7	263	2	A31299	chymotrypsin (EC 3
136	15	57.7	263	2	A21195	chymotrypsin (EC 3
137	15	57.7	264	2	T18136	unknown protein T3
138	15	57.7	264	2	F96807	hypothetical prote
139	15	57.7	273	2	A05113	translacton elonga
140	15	57.7	274	2	S64523	hypothetical prote
141	15	57.7	293	2	A85817	probable tail comp
142	15	57.7	293	2	E85718	hypothetical prote
143	15	57.7	293	2	T31840	hypothetical prote
144	15	57.7	294	2	E84640	CONSTANS-like B-po
145	15	57.7	297	2	H72779	hypothetical prote
146	15	57.7	307	2	T08125	CONSTANS protein h
147	15	57.7	307	2	T45213	superoxide dismuta
148	15	57.7	311	2	T15268	hypothetical prote
149	15	57.7	312	2	JCS522	uncoupling protein
150	15	57.7	312	2	T31834	hypothetical prote
151	15	57.7	323	2	S72915	hypothetical prote
152	15	57.7	326	2	T19996	hypothetical prote
153	15	57.7	337	2	T23794	hypothetical prote
154	15	57.7	344	2	S42383	hypothetical prote
155	15	57.7	345	2	A05279	surface antigen 51
156	15	57.7	363	2	JCS536	C-Fringe protein 1
157	15	57.7	368	2	S18758	hypothetical prote
158	15	57.7	379	2	B84073	sulfate adenyllyl tr
159	15	57.7	382	2	T27058	hypothetical prote
160	15	57.7	385	2	T25492	hypothetical prote
161	15	57.7	386	2	T33240	hypothetical prote
162	15	57.7	390	2	C95954	hypothetical prote
163	15	57.7	392	2	S66693	opioid receptor mu
164	15	57.7	395	1	A35629	mevalonate kinase
165	15	57.7	400	2	I56553	mu opiate receptor
166	15	57.7	410	2	T39115	formamide-like p
167	15	57.7	412	2	AF0668	transforming prote
168	15	57.7	423	1	TVFV2E	transforming prote
169	15	57.7	423	2	S38953	carboxypeptidase D
170	15	57.7	423	2	A41204	carboxypeptidase B
171	15	57.7	450	1	S13730	pmba protein - Bsc
172	15	57.7	450	2	AB1056	probable PmbA prot
173	15	57.7	450	2	D86121	maturation of anti
174	15	57.7	450	2	D91280	maturation of anti
175	15	57.7	454	2	T34297	hypothetical prote
176	15	57.7	459	2	I48854	gene murine tumour
177	15	57.7	461	1	A35356	tumor necrosis fac
178	15	57.7	470	2	A40697	63K sperm flagella
179	15	57.7	471	2	T39571	probable regulator
180	15	57.7	474	2	B38634	tumor necrosis fac
181	15	57.7	474	2	T43504	sp66 protein - fis
182	15	57.7	476	2	T29054	probable transmemb
183	15	57.7	482	2	S65766	G protein-coupled
184	15	57.7	488	2	T10794	phosphorin III -
185	15	57.7	493	2	S46826	hypothetical prote
186	15	57.7	493	2	S36931	chitinase (EC 3.2.
187	15	57.7	495	2	S31493	env polypeptid -
188	15	57.7	503	2	S59698	HSP1 protein - yea
189	15	57.7	504	2	C85485	probable carnitine
190	15	57.7	504	2	C90634	probable carnitine
191	15	57.7	504	2	H64724	probable carnitine
192	15	57.7	505	2	AD0511	probable carnitine
193	15	57.7	505	2	T19971	hypothetical prote
194	15	57.7	506	2	T19973	hypothetical prote
195	15	57.7	512	2	S28267	phosphorin I prec
196	15	57.7	514	2	T06585	ammonium transport
197	15	57.7	525	2	F84933	IMP cyclohydrolase
198	15	57.7	526	2	T07082	lycopen epsilon-c
199	15	57.7	529	2	S51477	extracellular glyc
200	15	57.7	529	2	T48253	myb-like protein -
201	15	57.7	533	2	T00742	ubiquitin-binding
202	15	57.7	543	2	S65462	glucose transport
203	15	57.7	545	2	A46281	tetrahydrofolylpol
204	15	57.7	546	2	T11217	reverse transcript
205	15	57.7	549	2	JN0553	triacylglycerol li
206	15	57.7	553	2	B88949	protein R0985.4 [i
207	15	57.7	553	2	T01479	hypothetical prote
208	15	57.7	561	2	T27318	hypothetical prote
209	15	57.7	562	1	RGBYS2	regulatory protein
210	15	57.7	568	2	T31692	hypothetical prote
211	15	57.7	571	2	T52325	pectinesterase (EC
212	15	57.7	576	2	H87414	TPR domain protein
213	15	57.7	584	2	S14952	pectinesterase hom
214	15	57.7	586	2	T52330	pectinesterase (EC
215	15	57.7	588	2	T02184	probable pectinest
216	15	57.7	595	2	T52327	pectinesterase (EC
217	15	57.7	595	2	T02183	probable pectinest
218	15	57.7	599	2	T10798	phosphorin-S - Vo
219	15	57.7	616	2	T29234	hypothetical prote
220	15	57.7	626	2	T27319	hypothetical prote
221	15	57.7	628	1	H64390	carbon-monoxide de
222	15	57.7	634	2	T27959	hypothetical prote
223	15	57.7	641	2	T05497	hypothetical prote
224	15	57.7	646	2	T35002	probable respirato
225	15	57.7	661	2	B71427	hypothetical prote
226	15	57.7	670	2	S36616	regulatory protein
227	15	57.7	676	2	T30480	envelope protein h
228	15	57.7	698	1	TFHUP	transferrin precu
229	15	57.7	698	2	D90771	hypothetical prote
230	15	57.7	698	2	H85633	hypothetical prote
231	15	57.7	698	2	B54839	yeaH protein precu
232	15	57.7	700	2	T32205	hypothetical prote
233	15	57.7	703	2	A45543	lactoferrin precu
234	15	57.7	713	2	A35502	major surface-labe
235	15	57.7	716	2	T09462	juvenile hormone r
236	15	57.7	717	2	S12100	transferrin precu
237	15	57.7	717	2	T25431	hypothetical prote
238	15	57.7	733	2	A35421	probable oxidoredu
239	15	57.7	738	2	T26421	hypothetical prote
240	15	57.7	738	2	B88593	hypothetical prote
241	15	57.7	744	2	S40092	protein R04H4.2b [
242	15	57.7	748	2	S57061	nuclear pore prote
243	15	57.7	748	2	B45046	basic juvenile hor
244	15	57.7	749	2	B35889	probable denhydrog
245	15	57.7	750	2	H91074	transcription regu
246	15	57.7	750	2	G85919	transcription regu
247	15	57.7	750	2	D65051	hyPF protein - Bsc
248	15	57.7	751	2	F87789	protein C34G6.2 [i

249	15	57.7	752	2	G85941	hypothetical prote	322	15	57.7	2090	2	T30075	hypothetical prote
250	15	57.7	752	2	C91096	hypothetical prote	323	15	57.7	2153	2	T30074	hypothetical prote
251	15	57.7	752	2	B65070	probable oxidoredu	324	15	57.7	2201	2	A54774	ATP binding caset
252	15	57.7	755	1	A44315	cartilage oligomer	325	15	57.7	2476	2	T34022	zonadhesin - pig
253	15	57.7	771	1	WZB852	gene 52 protein -	326	15	57.7	2531	2	T31070	notch homolog - se
254	15	57.7	793	1	KXMSF	furin (EC 3.4.21.7	327	15	57.7	2844	2	S28291	hypothetical prote
255	15	57.7	793	1	KXHP	furin (EC 3.4.21.7	328	15	57.7	2910	2	T42214	otogelin - mouse
256	15	57.7	794	1	KXHP	furin (EC 3.4.21.7	329	15	57.7	2918	2	A54105	fibillin-2 precur
257	15	57.7	797	1	I46044	furin (EC 3.4.21.7	330	15	57.7	3097	2	T00021	DN-cadherin - fru
258	15	57.7	806	2	T18840	hypothetical prote	331	15	57.7	3147	2	T21328	hypothetical prote
259	15	57.7	814	2	T49207	receptor kinase-11	332	15	57.7	3712	2	S18253	laminin alpha-1 ch
260	15	57.7	846	1	VCLJND	env polypeptide pr	333	15	57.7	4660	2	T42737	gp330 protein prec
261	15	57.7	847	1	S36337	histidine decarbox	334	15	57.7	4957	2	T03455	ALR protein - huma
262	15	57.7	849	2	T49342	fzr related protei	335	15	57.7	5262	2	T03454	subilosin A [vali
263	15	57.7	852	2	S41886	DNA repair protein	336	14	53.8	43	2	A69704	fulvicin C - Myxoc
264	15	57.7	853	2	S54384	envelope polypept	337	14	53.8	45	1	FRYZ	calixoxin - sea an
265	15	57.7	854	1	VCLJST	env polypeptide pr	338	14	53.8	46	2	A31863	vire7 protein - Ag
266	15	57.7	855	1	VCLJZR	env polypeptide pr	339	14	53.8	55	1	B7AGA6	lipoprotein R21 pr
267	15	57.7	856	1	A44963	gag polypeptide -	340	14	53.8	60	1	JN0750	lipoprotein R21 pr
268	15	57.7	863	2	A53034	hypothetical prote	341	14	53.8	60	2	H90831	metallothionein-2a
269	15	57.7	863	2	T27958	hypothetical prote	342	14	53.8	62	2	S54336	metallothionein-2c
270	15	57.7	873	2	B87049	conserved hypotet	343	14	53.8	62	2	S54335	hypothetical prote
271	15	57.7	888	1	GMLJHD	pol polypeptide -	344	14	53.8	62	2	T10302	hypothetical prote
272	15	57.7	890	2	T21000	hypothetical prote	345	14	53.8	65	2	S77379	hypothetical prote
273	15	57.7	891	2	E96590	hypothetical prote	346	14	53.8	73	2	F70641	hypothetical prote
274	15	57.7	895	2	B96775	aspartate kinase (	347	14	53.8	73	2	G97142	probable lipoprote
275	15	57.7	911	2	S46497	hypothetical prote	348	14	53.8	75	2	A10863	hypothetical prote
276	15	57.7	911	2	B86438	hypothetical prote	349	14	53.8	76	2	B91087	hypothetical prote
277	15	57.7	914	2	T17233	hypothetical prote	350	14	53.8	76	2	E65063	hypothetical prote
278	15	57.7	915	2	T21773	hypothetical prote	351	14	53.8	77	2	T10276	hypothetical prote
279	15	57.7	928	2	T20035	glycoprotein GP330	352	14	53.8	78	1	IKEC51	colicin V immunity
280	15	57.7	972	2	A30363	hypothetical prote	353	14	53.8	80	2	PM0044	estrogen receptor
281	15	57.7	975	2	T48107	hypothetical prote	354	14	53.8	83	2	H81188	hypothetical prote
282	15	57.7	978	2	H86319	hypothetical prote	355	14	53.8	85	2	B83008	hypothetical prote
283	15	57.7	982	1	VCLJLK	env polypeptide -	356	14	53.8	86	2	S78488	hypothetical prote
284	15	57.7	995	1	S50358	hypothetical prote	357	14	53.8	86	2	T07829	Ig kappa chain V r
285	15	57.7	1012	1	DJB86S	DNA-directed DNA p	358	14	53.8	88	2	JC1126	systemic acquired
286	15	57.7	1012	2	T44185	probable DNA-direc	359	14	53.8	88	2	S00036	major allergen cha
287	15	57.7	1012	2	T43998	DNA polymerase [lm	360	14	53.8	91	2	B87324	Clara cell 10K pro
288	15	57.7	1012	2	T41940	DNA polymerase - h	361	14	53.8	91	2	T37327	hypothetical prote
289	15	57.7	1013	2	QJ1920	DNA-directed DNA p	362	14	53.8	92	1	S36358	insulin homolog ce
290	15	57.7	1014	2	T30431	DNA-directed DNA p	363	14	53.8	92	1	UGMS	proteinase inhibit
291	15	57.7	1018	2	T19693	hypothetical prote	364	14	53.8	96	2	C71614	ribosomal protein
292	15	57.7	1047	2	S19508	MSH3 protein - yea	365	14	53.8	96	2	A36581	polychlorinated bi
293	15	57.7	1059	2	T22545	hypothetical prote	366	14	53.8	97	2	S72866	hypothetical prote
294	15	57.7	1063	2	T46284	hypothetical prote	367	14	53.8	98	2	A10334	conserved hypotet
295	15	57.7	1064	2	A40136	fibropellin 1a - s	368	14	53.8	99	2	UC2136	monocyte chemoatr
296	15	57.7	1083	2	S59780	hypothetical prote	369	14	53.8	99	2	S60230	glibetellin-regula
297	15	57.7	1095	2	T13964	probable histone d	370	14	53.8	100	2	S64316	hypothetical prote
298	15	57.7	1136	2	A48292	mucin, trachebron	371	14	53.8	102	2	F70534	hypothetical prote
299	15	57.7	1136	2	AH1227	different proteins	372	14	53.8	103	2	H84599	hypothetical prote
300	15	57.7	1193	2	T50729	magnesium-protopor	373	14	53.8	103	2	S70182	hypothetical prote
301	15	57.7	1224	2	T07446	DNA-directed RNA p	374	14	53.8	105	2	S42214	hydroxymethylgluta
302	15	57.7	1245	2	D86260	protein T12C24.22	375	14	53.8	105	2	S42212	hydroxymethylgluta
303	15	57.7	1274	2	T10729	transferrin-like p	376	14	53.8	105	2	S17345	hydroxymethylgluta
304	15	57.7	1274	2	T42017	cysteine rich prot	377	14	53.8	105	2	S42213	hydroxymethylgluta
305	15	57.7	1274	2	S06187	RNA2 polypeptide -	378	14	53.8	105	2	S42215	pectin esterase ho
306	15	57.7	1274	2	S06187	protein F1N21.4 [1	379	14	53.8	105	2	T14406	galactose-binding
307	15	57.7	1357	2	B96696	hypothetical prote	380	14	53.8	105	2	A37961	zinc finger protei
308	15	57.7	1391	2	T20406	protein-tyrosine k	381	14	53.8	105	4	B48827	hypothetical prote
309	15	57.7	1477	2	T18534	polypeptide - Afri	382	14	53.8	106	4	S57386	putidaredoxin [val
310	15	57.7	1524	2	T30337	hypothetical prote	383	14	53.8	107	1	T49527	hypothetical prote
311	15	57.7	1544	2	T04464	hypothetical prote	384	14	53.8	107	2	E72850	Acotf-5 protein
312	15	57.7	1584	2	T00026	brain-specific ang	385	14	53.8	109	2	T41895	orfs homolog orf13
313	15	57.7	1626	2	T26318	hypothetical prote	386	14	53.8	109	2	A13178	conserved hypotet
314	15	57.7	1687	2	T30176	BGF repeat: transme	387	14	53.8	109	2	S64475	hypothetical prote
315	15	57.7	1722	2	E89753	protein F1IC7.4 [1	388	14	53.8	111	2	S62929	hypothetical prote
316	15	57.7	1743	2	T26859	hypothetical prote	389	14	53.8	112	2	S07216	gonadotropin I bet
317	15	57.7	1820	2	A55494	latent transforin	390	14	53.8	113	2	T20299	hypothetical prote
318	15	57.7	1827	2	A35694	cut1 protein - fis	391	14	53.8	113	2	S15197	hypA protein - Esc
319	15	57.7	1828	2	T41455	cut1 protein - fis	392	14	53.8	116	1	S15197	pleiotropic effec
320	15	57.7	1863	1	A58881	genome polypeptide	393	14	53.8	116	2	F85921	hypA protein (limpo
321	15	57.7	1866	1	GNWE2C		394	14	53.8	116	2	F91076	

395	14	53.8	116	2	A54598	468	14	53.8	154	2	S12542	gene X protein - h
396	14	53.8	116	2	D95280	469	14	53.8	154	2	S33687	gene X protein - h
397	14	53.8	116	2	F81285	470	14	53.8	155	2	D72761	hypothetical prote
398	14	53.8	118	1	PSSNKL	471	14	53.8	156	2	I67751	dopamine receptor
399	14	53.8	118	1	PSSNKL	472	14	53.8	156	2	B83185	probable transcrip
400	14	53.8	118	2	D34860	473	14	53.8	156	2	B82073	sigma-E factor reg
401	14	53.8	118	2	E34860	474	14	53.8	159	1	I83299	sigma-E factor reg
402	14	53.8	118	2	G34860	475	14	53.8	159	2	AF0829	sigma-E factor reg
403	14	53.8	118	2	F34860	476	14	53.8	159	2	B85903	sigma-E factor, ne
404	14	53.8	118	2	H34860	477	14	53.8	159	2	D91058	sigma-E factor reg
405	14	53.8	118	2	AB0847	478	14	53.8	159	2	T02999	hypothetical prote
406	14	53.8	119	2	AC3199	479	14	53.8	160	2	T28953	globin chain al pr
407	14	53.8	120	2	F70424	480	14	53.8	160	2	JC7606	hypothetical prote
408	14	53.8	121	2	H71351	481	14	53.8	161	2	T21405	protein T21E12.3 l
409	14	53.8	123	2	T04635	482	14	53.8	162	2	E87755	probable deoxycyri
410	14	53.8	124	2	E70754	483	14	53.8	166	2	E71342	hypothetical prote
411	14	53.8	125	2	AF2154	484	14	53.8	167	2	D81049	hypothetical prote
412	14	53.8	127	2	F70910	485	14	53.8	171	2	S09903	hypothetical prote
413	14	53.8	128	2	D69326	486	14	53.8	171	2	S09759	probable acetyltra
414	14	53.8	129	2	AB3212	487	14	53.8	171	2	B83360	hypothetical integ
415	14	53.8	131	1	BGB02	488	14	53.8	172	2	B81828	invasion protein b
416	14	53.8	133	2	AB1108	489	14	53.8	173	2	AB3450	probable membrane
417	14	53.8	137	2	B36179	490	14	53.8	174	2	AB1923	small hydrophobic
418	14	53.8	137	2	I51231	491	14	53.8	174	1	JQ1625	hypothetical prote
419	14	53.8	137	2	S34349	492	14	53.8	176	2	T28762	conserved hypotet
420	14	53.8	137	2	S55364	493	14	53.8	177	2	B81085	protective antigen
421	14	53.8	138	2	T21792	494	14	53.8	177	2	S33166	conserved hypotet
422	14	53.8	139	2	S69458	495	14	53.8	177	2	H95981	conserved hypotet
423	14	53.8	139	2	PC4217	496	14	53.8	178	2	AB4257	hypothetical prote
424	14	53.8	139	2	T12618	497	14	53.8	178	2	S64257	EST3 protein - yea
425	14	53.8	140	2	T04904	498	14	53.8	181	2	T78572	hypothetical prote
426	14	53.8	141	2	A85994	499	14	53.8	182	2	T17625	hypothetical prote
427	14	53.8	141	2	AE1009	500	14	53.8	184	2	D83234	hypothetical prote
428	14	53.8	141	2	E91148	501	14	53.8	185	2	T29777	21K protein precu
429	14	53.8	141	2	I67892	502	14	53.8	187	2	T09390	hypothetical prote
430	14	53.8	141	2	T09251	503	14	53.8	189	2	T28554	Al2u protein - var
431	14	53.8	142	2	T51316	504	14	53.8	189	2	D36849	Al2u protein - var
432	14	53.8	142	2	T28293	505	14	53.8	189	2	B72165	Al3u protein - var
433	14	53.8	142	2	E84040	506	14	53.8	189	2	G95514	core protein MGC-2
434	14	53.8	142	2	JS0510	507	14	53.8	189	2	JK0235	hypothetical 21.5K
435	14	53.8	142	2	T03976	508	14	53.8	191	2	JK1999	hypothetical prote
436	14	53.8	142	4	S13768	509	14	53.8	191	2	T32278	Al2u protein - vac
437	14	53.8	143	2	H83935	510	14	53.8	192	2	E42518	hypothetical prote
438	14	53.8	143	2	T00293	511	14	53.8	193	2	D83296	hypothetical prote
439	14	53.8	143	2	T27692	512	14	53.8	196	2	T25943	hypothetical prote
440	14	53.8	144	2	S01391	513	14	53.8	197	2	E85068	hypothetical prote
441	14	53.8	145	2	T05641	514	14	53.8	197	2	T26829	hypothetical prote
442	14	53.8	146	2	T25356	515	14	53.8	197	2	C88809	protein Y43C58.1 l
443	14	53.8	146	2	C86187	516	14	53.8	197	2	S72709	Leb1170 C3 229 pr
444	14	53.8	146	2	AD3427	517	14	53.8	198	2	A36646	ORF1 protein - equ
445	14	53.8	147	2	S42552	518	14	53.8	198	2	B36646	hypothetical prote
446	14	53.8	148	2	T30627	519	14	53.8	200	2	B83642	resolvasin - Klebs
447	14	53.8	148	2	T48981	520	14	53.8	201	2	S38652	hypothetical prote
448	14	53.8	148	2	B72513	521	14	53.8	201	2	D82779	hypothetical prote
449	14	53.8	148	2	S11441	522	14	53.8	203	2	E70150	hypothetical prote
450	14	53.8	150	2	B83550	523	14	53.8	204	2	T02386	hypothetical prote
451	14	53.8	151	2	T09256	524	14	53.8	205	2	T02385	hypothetical prote
452	14	53.8	152	1	KRSHHC	525	14	53.8	205	2	T04918	hypothetical prote
453	14	53.8	152	2	I47111	526	14	53.8	207	2	T22436	hypothetical prote
454	14	53.8	152	2	I47109	527	14	53.8	207	2	S70533	bBk2.10 protein pr
455	14	53.8	152	2	I47108	528	14	53.8	208	2	B75458	hypothetical prote
456	14	53.8	152	2	I47112	529	14	53.8	208	2	P00278	hypothetical prote
457	14	53.8	152	2	A84503	530	14	53.8	209	2	C70857	probable Igpa prot
458	14	53.8	152	2	T09253	531	14	53.8	211	2	S25938	env protein - huma
459	14	53.8	152	2	D89753	532	14	53.8	211	2	T18799	hypothetical prote
460	14	53.8	153	2	F75614	533	14	53.8	212	2	S09623	agglutinin isolat
461	14	53.8	153	2	T32950	534	14	53.8	213	1	AEW72	hypothetical prote
462	14	53.8	154	1	QQVILAW	535	14	53.8	213	2	T27841	conserved hypotet
463	14	53.8	154	1	QQVILCP	536	14	53.8	213	2	B82698	hemiferrin - bovin
464	14	53.8	154	1	QQVILKS	537	14	53.8	215	2	A60166	hemiferrin - bovin
465	14	53.8	154	2	S47408	538	14	53.8	215	2	A71541	hemiferrin - rat
466	14	53.8	154	2	S20756	539	14	53.8	216	1	A39684	protein W10G11.13
467	14	53.8	154	2	S67503	540	14	53.8	216	2	H88102	



541	14	53.8	216	2	S60048	chlorophyll a/c-bi	614	14	53.8	273	2	G83403	conserved hypochet
542	14	53.8	217	2	T25618	hypothetical prote	615	14	53.8	274	2	UC4171	tryptase (EC 3.4.2
543	14	53.8	218	2	D69428	hypothetical prote	616	14	53.8	275	2	AE3112	transcription regu
544	14	53.8	219	2	T34373	hypothetical prote	617	14	53.8	276	2	H98174	transcription regu
545	14	53.8	220	2	S25939	env protein - huma	618	14	53.8	277	2	A32410	tryptase (EC 3.4.2
546	14	53.8	221	2	D83390	carbonate dehydrat	619	14	53.8	278	2	G83522	conserved hypochet
547	14	53.8	222	1	MMVZB4	antithrombin-III h	620	14	53.8	279	2	T29952	hypothetical prote
548	14	53.8	223	2	AW5172	hypothetical prote	621	14	53.8	280	2	S71222	xyloglucan endo-1,
549	14	53.8	224	2	P89567	protein T08A9.3 [i	622	14	53.8	281	2	D97185	dtDP-4-dehydroxam
550	14	53.8	225	2	T37974	probable peroxisom	623	14	53.8	282	2	G01884	LIM protein FHL-1,
551	14	53.8	226	2	T01325	hypothetical prote	624	14	53.8	283	2	F95998	probable sugar nuc
552	14	53.8	227	2	C86563	CT465 hypothetical	625	14	53.8	284	2	AE2248	phosphomethylpyrim
553	14	53.8	228	2	C72062	conserved hypochet	626	14	53.8	285	2	C86221	hypothetical prote
554	14	53.8	229	2	T27843	hypothetical prote	627	14	53.8	286	2	T15556	hypothetical prote
555	14	53.8	230	2	T27019	hypothetical prote	628	14	53.8	287	2	T15640	hypothetical prote
556	14	53.8	231	2	T27840	hypothetical prote	629	14	53.8	288	2	T18969	5'-methylchideno
557	14	53.8	232	2	B82544	hypothetical prote	630	14	53.8	289	2	T21622	hypothetical prote
558	14	53.8	233	2	S28703	hypothetical prote	631	14	53.8	290	2	H97220	transcriptional reg
559	14	53.8	234	2	F64642	hypothetical prote	632	14	53.8	291	2	S04278	xyloglucan endo-1,
560	14	53.8	235	2	T24669	hypothetical prote	633	14	53.8	292	2	T06200	hypoxanthine phosph
561	14	53.8	236	2	E71325	probable V-type AT	634	14	53.8	293	1	IDECRP	replication initia
562	14	53.8	237	2	T32843	hypothetical prote	635	14	53.8	294	2	I64780	replication-associ
563	14	53.8	238	2	T28070	hypothetical prote	636	14	53.8	295	2	T00306	replication initia
564	14	53.8	239	2	T04166	thiamatin-like pro	637	14	53.8	296	2	B88546	protein R107.2 [im
565	14	53.8	240	2	T40568	hypothetical prote	638	14	53.8	297	2	T24665	hypoxanthine phosph
566	14	53.8	241	2	B81345	probable periplasm	639	14	53.8	298	2	S09614	hypothetical prote
567	14	53.8	242	2	T29854	hypothetical prote	640	14	53.8	299	2	S30872	hypothetical prote
568	14	53.8	243	2	T29854	hypothetical prote	641	14	53.8	300	2	B16570	ethanolamine ammon
569	14	53.8	244	2	A33329	testis-specific pr	642	14	53.8	301	2	T23354	hypothetical prote
570	14	53.8	245	2	JB0204	testicular protein	643	14	53.8	302	2	T31494	hypothetical prote
571	14	53.8	246	1	UC2018	phatucatin - mouse	644	14	53.8	303	2	A85059	hypothetical prote
572	14	53.8	247	2	T37493	phatucatin oxidore	645	14	53.8	304	2	T02546	hypothetical prote
573	14	53.8	248	2	S52096	phosducin - cat	646	14	53.8	305	2	T21790	hypothetical prote
574	14	53.8	249	2	AP0412	sugar fermentation	647	14	53.8	306	2	T34688	probable lipoprote
575	14	53.8	250	2	T24565	hypothetical prote	648	14	53.8	307	2	C71321	hypothetical prote
576	14	53.8	251	2	A35422	phosducin, retinal	649	14	53.8	308	2	T10104	maturase-like prot
577	14	53.8	252	1	T43324	synaptogyrin homol	650	14	53.8	309	2	T19669	hypothetical prote
578	14	53.8	253	2	S49323	chymotrypsin (EC 3	651	14	53.8	310	2	G88783	protein C33A12.16
579	14	53.8	254	2	T21786	hypothetical prote	652	14	53.8	311	2	B49539	xyloglucan endo-1,
580	14	53.8	255	2	T22769	hypothetical prote	653	14	53.8	312	2	A48844	TGF alpha-like pro
581	14	53.8	256	2	T31837	hypothetical prote	654	14	53.8	313	2	G65018	ethanolamine ammon
582	14	53.8	257	2	T21785	hypothetical prote	655	14	53.8	314	2	B85887	ethanolamine ammon
583	14	53.8	258	2	T31836	hypothetical prote	656	14	53.8	315	2	G91042	ethanolamine ammon
584	14	53.8	259	2	T31835	hypothetical prote	657	14	53.8	316	2	C90740	probable transmemb
585	14	53.8	260	2	T14407	pectin esterase ho	658	14	53.8	317	2	A46017	zinc finger protei
586	14	53.8	261	2	A35523	hypothetical prote	659	14	53.8	318	2	B85590	probable transmemb
587	14	53.8	262	2	B35685	GRBAG protein 2.1	660	14	53.8	319	2	B64818	probable membrane
588	14	53.8	263	2	AB0213	conserved hypochet	661	14	53.8	320	2	H87511	hypothetical prote
589	14	53.8	264	2	S61178	hypothetical prote	662	14	53.8	321	2	T18660	hypothetical prote
590	14	53.8	265	2	T04059	hypothetical prote	663	14	53.8	322	2	A98294	hypothetical prote
591	14	53.8	266	2	B89130	protein F52E1.2 [i	664	14	53.8	323	2	AH2989	permease (imported
592	14	53.8	267	2	AB1304	dihydroorotate deh	665	14	53.8	324	2	S06267	surface antigen H
593	14	53.8	268	2	AB1676	dihydroorotate deh	666	14	53.8	325	2	AF0813	ethanolamine ammon
594	14	53.8	269	2	F82304	sugar fermentation	667	14	53.8	326	2	T43989	minor capsid proteo
595	14	53.8	270	2	T28946	hypothetical prote	668	14	53.8	327	2	C88970	transcription regu
596	14	53.8	271	2	T47491	hypothetical prote	669	14	53.8	328	2	F82287	transcription regu
597	14	53.8	272	2	B82249	uridine phosphoryl	670	14	53.8	329	2	B84325	tRNA-pseudouridine
598	14	53.8	273	2	259	insulin-like grow	671	14	53.8	330	2	D65054	hypothetical prote
599	14	53.8	274	2	10H01	hypothetical prote	672	14	53.8	331	2	H84731	hypothetical prote
600	14	53.8	275	2	T27783	hypothetical prote	673	14	53.8	332	2	UC2071	chitinase (EC 3.2.
601	14	53.8	276	2	E84427	hypothetical prote	674	14	53.8	333	2	A72596	hypothetical prote
602	14	53.8	277	2	T05084	hypothetical prote	675	14	53.8	334	2	T28005	hypothetical prote
603	14	53.8	278	2	S44668	ZK370.7 protein -	676	14	53.8	335	2	S32834	methylviologen-red
604	14	53.8	279	2	T09324	capsid protein - h	677	14	53.8	336	2	T41985	hypothetical prote
605	14	53.8	280	2	AE2274	D-2-chlorotriptoni	678	14	53.8	337	2	A28446	transferrin - mous
606	14	53.8	281	2	S62363	conserved hypochet	679	14	53.8	338	2	UC7873	L-thiamose-binding
607	14	53.8	282	2	C81651	polymorphic outer	680	14	53.8	339	2	T09443	zinc metalloprotei
608	14	53.8	283	2	H86548	complement factor	681	14	53.8	340	2	B82276	hypothetical prote
609	14	53.8	284	2	137278	unknown protein F5	682	14	53.8	341	2	T17969	hypothetical prote
610	14	53.8	285	2	C96731	protein F13E6.3 [i	683	14	53.8	342	2	T45061	hypothetical prote
611	14	53.8	286	2	P89632	hypothetical prote	684	14	53.8	343	2	T28879	hypothetical prote
612	14	53.8	287	2	T40238	hypothetical prote	685	14	53.8	344	2	A33985	wound-inducible ch
613	14	53.8	288	2	T15351	hypothetical prote	686	14	53.8	345	2	T04403	probable chitinase

687	14	53.8	319	2	H83253	conserved hypothet	760	14	53.8	369	2	S60186	peridin-chloroph
688	14	53.8	319	2	B86395	probable origin re	761	14	53.8	370	2	S15013	wnt-1 protein - ze
689	14	53.8	320	2	S36670	chitinase (EC 3.2.	762	14	53.8	372	1	W2ML21	E2 protein - human
690	14	53.8	320	2	T28992	hypothetical prote	763	14	53.8	372	2	T31060	hypothetical prote
691	14	53.8	322	1	A29928	membrane-associate	764	14	53.8	372	2	T45410	hypothetical prote
692	14	53.8	322	2	H87131	probable transcrip	765	14	53.8	372	2	T29359	hypothetical prote
693	14	53.8	323	2	G87358	diene lactone hydro	766	14	53.8	373	2	A70856	probable 1ppz prot
694	14	53.8	323	2	T33279	hypothetical prote	767	14	53.8	373	2	T47488	hypothetical prote
695	14	53.8	323	2	A47172	transforming growt	768	14	53.8	375	2	G02300	pancreatic polypep
696	14	53.8	324	2	JC2395	Fas antigen precu	769	14	53.8	375	2	T3182	neuropeptide Y/pep
697	14	53.8	325	2	I65354	stem cell antigen	770	14	53.8	375	2	S63685	neuropeptide Y rec
698	14	53.8	327	2	F96814	hypothetical prote	771	14	53.8	375	2	S52143	amac protein - Erw
699	14	53.8	328	2	G89152	protein C24B5.5 f1	772	14	53.8	376	2	A96720	hypothetical prote
700	14	53.8	329	2	T22636	hypothetical prote	773	14	53.8	378	2	E64044	rod shape-determ
701	14	53.8	329	2	T16619	hypothetical prote	774	14	53.8	379	2	G82380	hypothetical prote
702	14	53.8	331	2	T06227	peroxidase (EC 1.1	775	14	53.8	380	2	C86754	Similar to part of
703	14	53.8	331	2	C71468	probable glycerol-	776	14	53.8	381	2	G83835	sulfate adenyllyl tr
704	14	53.8	331	2	B83486	cobalamin biosynth	777	14	53.8	382	2	T27076	hypothetical prote
705	14	53.8	332	2	T04484	probable chitinase	778	14	53.8	382	2	B46133	neuropeptide Y/pep
706	14	53.8	332	2	T21458	hypothetical prote	779	14	53.8	382	2	S27388	neuropeptide Y rec
707	14	53.8	332	2	A36663	nautilus protein -	780	14	53.8	382	2	I52565	stem cell antigen
708	14	53.8	333	2	D88970	protein F15E1.11	781	14	53.8	382	2	E91083	hypothetical membr
709	14	53.8	333	2	T34498	hypothetical prote	782	14	53.8	382	2	P85928	glycoprotein precu
710	14	53.8	334	2	B81272	probable sugar nuc	783	14	53.8	383	1	VGBBKG	neuropeptide Y/pep
711	14	53.8	336	2	T15727	hypothetical prote	784	14	53.8	384	2	A45490	retrovirus-related
712	14	53.8	340	2	JC7695	G protein-coupled	785	14	53.8	384	2	S64735	probable exopolyp
713	14	53.8	341	1	WMVZHI	antithrombin-III h	786	14	53.8	384	2	T38544	dopamine D1-like r
714	14	53.8	342	1	H89980	conserved hypothet	787	14	53.8	385	2	S68780	homeotic protein d
715	14	53.8	344	2	E72173	B13R protein - vari	788	14	53.8	385	2	S53718	preadipocyte facto
716	14	53.8	344	2	S46871	hypothetical prote	789	14	53.8	385	2	A54785	dopamine receptor
717	14	53.8	344	2	T28611	probable hydrolase	790	14	53.8	386	2	S72168	hypothetical prote
718	14	53.8	344	2	A70771	antithrombin-III h	791	14	53.8	386	2	T12527	hypothetical prote
719	14	53.8	345	1	WMVZMW	apolipoprotein H p	792	14	53.8	387	2	B71611	hypothetical prote
720	14	53.8	345	1	UN0465	hypothetical prote	793	14	53.8	388	2	S51316	prostaglandin E re
721	14	53.8	346	2	T32018	siatoglycoproteina	794	14	53.8	388	2	I38750	hypothetical 43. 6K
722	14	53.8	346	2	A71820	hypothetical prote	795	14	53.8	390	2	S56560	hypothetical prote
723	14	53.8	346	2	T45748	hypothetical prote	796	14	53.8	390	2	B91291	hypothetical prote
724	14	53.8	346	2	T47411	hypothetical prote	797	14	53.8	390	2	D6132	keratin 2, type I,
725	14	53.8	347	2	S60428	hypothetical prote	798	14	53.8	392	2	A60777	hypothetical prote
726	14	53.8	347	2	UQ2359	wheat albumin ind	799	14	53.8	392	2	T33505	hypothetical prote
727	14	53.8	348	2	S64041	probable membrane	800	14	53.8	392	2	S58157	mevalonate kinase
728	14	53.8	348	2	S75548	siatoglycoproteina	801	14	53.8	395	2	A42919	conserved hypothet
729	14	53.8	349	2	S19873	single-stranded DN	802	14	53.8	396	1	T08345	hypothetical prote
730	14	53.8	349	2	AP3401	enoyl-CoA hydratase	803	14	53.8	397	2	AH0482	conserved hypothet
731	14	53.8	349	2	S12863	G protein-coupled	804	14	53.8	398	2	T19421	hypothetical prote
732	14	53.8	352	2	C96643	hypothetical prote	805	14	53.8	398	2	S32804	beta-3-adrenergic
733	14	53.8	352	2	S77448	hypothetical prote	806	14	53.8	400	2	A41679	beta-3-adrenergic
734	14	53.8	354	2	T26599	hypothetical prote	807	14	53.8	400	2	T49480	beta-3-adrenergic
735	14	53.8	354	2	G86276	hypothetical prote	808	14	53.8	400	2	A53281	conserved hypothet
736	14	53.8	356	2	A25918	thrombomodulin - b	809	14	53.8	400	2	D89810	hypothetical prote
737	14	53.8	356	2	T18590	hypothetical prote	810	14	53.8	401	2	E96570	hypothetical prote
738	14	53.8	357	2	A71295	probable alanine r	811	14	53.8	402	2	T39591	cyclin - fission Y
739	14	53.8	357	2	A97697	hypothetical prote	812	14	53.8	403	2	C82228	probable trypsin Y
740	14	53.8	358	2	D84901	hypothetical prote	813	14	53.8	403	2	A31847	hypothetical prote
741	14	53.8	358	2	T23802	hypothetical prote	814	14	53.8	403	2	T49480	hypothetical prote
742	14	53.8	359	2	AD3520	galactoside transp	815	14	53.8	405	2	S65459	beta-3-adrenergic
743	14	53.8	361	2	T25305	hypothetical prote	816	14	53.8	405	2	B36340	alpha(1, 3)-fucosyl
744	14	53.8	361	2	C65132	hypothetical 38.6	817	14	53.8	406	2	A43850	glutamate N-acetyl
745	14	53.8	362	1	LPBCRA	rate lipoprotein A	818	14	53.8	406	2	C81017	amino-acid N-acety
746	14	53.8	362	2	C85563	a minor lipoprotei	819	14	53.8	407	1	T19895	hypothetical prote
747	14	53.8	362	2	G90712	hypothetical prote	820	14	53.8	408	1	ORHUBE	beta-3-adrenergic
748	14	53.8	362	2	T20222	dopamine D1 recept	821	14	53.8	408	1	MMNVIA	immediate-early pr
749	14	53.8	363	2	I50475	triacylglycerol 11	822	14	53.8	408	2	B83333	probable MFS trans
750	14	53.8	363	2	A34959	probable acyl-CoA	823	14	53.8	408	2	A72869	early gene transac
751	14	53.8	363	2	A81134	acyl-CoA dehydroge	824	14	53.8	410	2	B84144	hypothetical prote
752	14	53.8	363	2	C81888	probable acyl-CoA	825	14	53.8	412	2	T26056	hypothetical prote
753	14	53.8	364	2	JC2115	prostaglandin E re	826	14	53.8	412	2	ORHUB2	beta-2-adrenergic
754	14	53.8	364	2	S65009	collagen alpha 2(I	827	14	53.8	413	1	UC5527	extracellular sucr
755	14	53.8	365	2	S10847	neuropeptide Y/pep	828	14	53.8	413	2	UC5520	beta-fructofuranos
756	14	53.8	366	1	S71152	GTP-binding protei	829	14	53.8	414	1	QRHUB3	beta-3-adrenergic
757	14	53.8	367	2	I51426	pectate lyase (EC	830	14	53.8	415	2	I53040	beta-2 adrenergic
758	14	53.8	367	2	S15716	cinnamoyl CoA redu	831	14	53.8	415	2	A64299	coenzyme F420 hydr
759	14	53.8	368	2	T46643		832	14	53.8	415	2		

833	14	53.8	415	2	S37690	DNA-binding protei	906	14	53.8	456	2	C71402	probable glucosylt
834	14	53.8	415	2	S34140	DNA-binding protei	907	14	53.8	457	2	I51660	dopamine D1B recep
835	14	53.8	415	2	T26057	hypothetical prote	908	14	53.8	457	2	T32244	hypothetical prote
836	14	53.8	416	2	T45051	hypothetical prote	909	14	53.8	459	2	A25928	cellulase (EC 3.2.
837	14	53.8	418	1	QRHYB2	beta-2-adrenergic	910	14	53.8	459	2	A56849	dopamine recep-t-
838	14	53.8	418	2	G02953	beta-3-adrenergic	911	14	53.8	460	2	T39462	argininosuccinate
839	14	53.8	418	2	S00260	beta-2-adrenergic	912	14	53.8	460	2	T48137	copper amine oxida
840	14	53.8	418	2	S10855	beta-2-adrenergic	913	14	53.8	461	2	H64087	L-tyrosyl-tRNAse se
841	14	53.8	419	2	S34421	GTP-binding regula	914	14	53.8	462	2	A10492	L-tyrosyl-tRNAse se
842	14	53.8	420	2	S06579	finger protein (c1	915	14	53.8	463	2	A48375	cellulase (EC 3.2.
843	14	53.8	420	2	A53531	oncofetal trophobl	916	14	53.8	463	2	B56849	dopamine recep-t-
844	14	53.8	421	2	T51055	hypothetical prote	917	14	53.8	464	2	UC7143	endoglucanase I -
845	14	53.8	423	2	A49772	cuticle protein cu	918	14	53.8	464	2	S12591	beta-1-adrenergic
846	14	53.8	424	2	T20032	hypothetical prote	919	14	53.8	464	2	T34332	hypothetical prote
847	14	53.8	424	2	T01383	GTPase-activating	920	14	53.8	465	2	I51661	dopamine D1C recep
848	14	53.8	424	2	T33552	hypothetical prote	921	14	53.8	466	2	S36794	beta-1-adrenergic
849	14	53.8	426	2	D88103	protein W10G1.6 f	922	14	53.8	467	2	T31522	hypothetical prote
850	14	53.8	427	2	S17148	alpha-chromidin rec	923	14	53.8	468	2	S02514	n1B protein - Kie
851	14	53.8	428	2	A55044	beta-4C-adrenergic	924	14	53.8	468	2	T22598	hypothetical prote
852	14	53.8	428	2	S03677	finger protein (c1	925	14	53.8	468	2	C95405	probable sulfite o
853	14	53.8	428	2	T48008	hypothetical prote	926	14	53.8	470	2	S26841	retrovirus-related
854	14	53.8	429	2	S09141	Ndt4, intron 2 prot	927	14	53.8	470	2	AF2828	FAD dependent oxid
855	14	53.8	430	2	S58169	porphobilinogen sy	928	14	53.8	470	2	T49175	hypothetical prote
856	14	53.8	430	2	T16079	hypothetical prote	929	14	53.8	470	2	T45816	hypothetical prote
857	14	53.8	430	2	E90869	hypothetical prote	930	14	53.8	471	2	S18446	variant surface gl
858	14	53.8	430	2	T25962	hypothetical prote	931	14	53.8	471	2	S76290	hypothetical prote
859	14	53.8	430	2	D85749	hypothetical prote	932	14	53.8	472	2	S72909	uroporphyrin-III C
860	14	53.8	430	2	H64883	conserved hypothet	933	14	53.8	474	2	S50957	probable membrane
861	14	53.8	431	2	T18753	hypothetical prote	934	14	53.8	475	2	P70972	probable amidase -
862	14	53.8	431	2	G72285	hydrid cluster [4f	935	14	53.8	475	2	A41271	dopamine recep-tor
863	14	53.8	431	2	P85097	hypothetical prote	936	14	53.8	477	1	QRHUB5	dopamine recep-tor
864	14	53.8	432	2	A43448	thrombin receptor	937	14	53.8	477	1	QRHUB5	beta-1-adrenergic
865	14	53.8	433	2	T21232	hypothetical prote	938	14	53.8	477	2	S71400	estrogen receptor
866	14	53.8	433	2	P96974	aspartyl aminopept	939	14	53.8	477	2	D97606	glycolate oxidase
867	14	53.8	434	2	G97610	hypothetical metab	940	14	53.8	478	2	G84576	probable beta-1,3-
868	14	53.8	434	2	AC2833	hypothetical prote	941	14	53.8	480	2	A63491	3-phosphoshikimate
869	14	53.8	434	2	S58476	non-replicative mo	942	14	53.8	480	2	I53053	beta 1 adrenergic
870	14	53.8	434	2	T47575	hypothetical prote	943	14	53.8	480	2	T00971	probable disease r
871	14	53.8	434	2	A38282	p58 galactosyltran	944	14	53.8	480	2	T20168	hypothetical prote
872	14	53.8	435	2	T16470	hypothetical prote	945	14	53.8	482	2	T23608	hypothetical prote
873	14	53.8	435	2	P86601	protoporphyrinogen	946	14	53.8	482	2	JC5092	beta-selectin - pig
874	14	53.8	435	2	D72023	protoporphyrinogen	947	14	53.8	483	2	A25896	beta-adrenergic re
875	14	53.8	435	2	A81518	protoporphyrinogen	948	14	53.8	483	2	S52567	retrovirus-related
876	14	53.8	436	2	AB2417	hypothetical prote	949	14	53.8	484	1	A46217	GPI-anchor biosynt
877	14	53.8	437	2	S34959	MAH2 dehydrogenas	950	14	53.8	484	2	S36339	phorboloxin II - V
878	14	53.8	437	2	T48035	hypothetical prote	951	14	53.8	484	2	T10668	hypothetical prote
879	14	53.8	439	2	A36385	surface antigen se	952	14	53.8	485	1	A55731	GPI-anchor biosynt
880	14	53.8	439	2	T09214	choline monooxygen	953	14	53.8	485	2	I52484	gene Pig-A protein
881	14	53.8	440	2	B71858	adenylosuccinate 1	954	14	53.8	485	2	T49237	hypothetical prote
882	14	53.8	440	2	H64658	adenylosuccinate 1	955	14	53.8	486	1	JC7241	phosphoprotein pho
883	14	53.8	440	2	T32190	hypothetical prote	956	14	53.8	486	1	COBYW2	WH12 protein - Yea
884	14	53.8	442	2	T10718	anthranilate N-ben	957	14	53.8	486	2	B55886	dopamine recep-tor
885	14	53.8	442	2	C86471	51.1K hypothetical	958	14	53.8	487	1	DYRTD1	dopamine recep-tor
886	14	53.8	444	2	C55886	dopamine recep-tor	959	14	53.8	487	2	T49424	hypothetical prote
887	14	53.8	444	2	T20374	hypothetical prote	960	14	53.8	487	2	T05127	hypothetical prote
888	14	53.8	445	2	T10717	anthranilate N-ben	961	14	53.8	488	1	URBYM	funarate hydratase
889	14	53.8	445	2	T10719	anthranilate N-ben	962	14	53.8	489	2	A49388	catalase (EC 1.11.
890	14	53.8	446	1	DYHDD1	dopamine recep-tor	963	14	53.8	489	2	D87551	glutamy1-tRNA(Gln)
891	14	53.8	446	2	T47217	dopamine recep-tor	964	14	53.8	491	2	JC2497	cyclin E - mouse
892	14	53.8	446	2	PC4393	Iva2 protein - Ovl	965	14	53.8	493	2	B77724	glutamy1-tRNA amid
893	14	53.8	446	2	T10711	anthranilate N-ben	966	14	53.8	493	2	E71725	glutamy1-tRNA amid
894	14	53.8	446	2	B70543	hypothetical prote	967	14	53.8	493	2	T25357	hypothetical prote
895	14	53.8	446	2	T14542	choline monooxygen	968	14	53.8	496	2	S37379	catalase (EC 1.11.
896	14	53.8	447	2	S25817	hypothetical prote	969	14	53.8	497	2	S47890	catalase (EC 1.11.
897	14	53.8	448	2	B85077	hypothetical prote	970	14	53.8	498	2	D86298	hypothetical prote
898	14	53.8	449	2	E36676	hypothetical prote	971	14	53.8	499	2	S63465	dihydrofolipomade d
899	14	53.8	450	2	F82659	UDP-glucose dehydr	972	14	53.8	500	1	JE0258	cyclochrome P450 2D
900	14	53.8	450	2	A55886	dopamine recep-tor	973	14	53.8	504	2	JQ2151	macron-type locus
901	14	53.8	450	2	T47450	n-acetylglucosamin	974	14	53.8	506	2	A40218	envelop glycoprote
902	14	53.8	450	2	G66670	hypothetical prote	975	14	53.8	506	2	T01292	hypothetical prote
903	14	53.8	451	2	I51659	dopamine D1A recep	976	14	53.8	508	2	D85056	probable aspartic
904	14	53.8	453	2	G86396	protein TN9.2 [1m	977	14	53.8	509	1	VCFPGT	retrovirus-related
905	14	53.8	453	2	T09745	myd-related protei	978	14	53.8	509	2	T00627	hypothetical prote

979 14 53.8 510 1 S15620 l1 protein - human  
980 14 53.8 510 1 S15627 l1 protein - human  
981 14 53.8 511 2 S44275 dopamine receptor  
982 14 53.8 511 2 UC7682 spermatogenesis as  
983 14 53.8 513 2 T38866 conserved hypothet  
984 14 53.8 515 2 D73119 probable glucose-6  
985 14 53.8 516 2 T33269 hypothetical prote  
986 14 53.8 516 2 S40996 hypothetical prote  
987 14 53.8 517 2 A71707 lipopolysaccharide  
988 14 53.8 519 2 S78089 g-protein signalin  
989 14 53.8 520 2 T23545 hypothetical prote  
990 14 53.8 522 2 S71821 probable interleuk  
991 14 53.8 522 1 S61713 carboxypeptidase C  
992 14 53.8 524 1 P1WL58 l1 protein - human  
993 14 53.8 527 1 SAHUP saposin precursor  
994 14 53.8 528 2 B88989 protein F02C9.2 l1  
995 14 53.8 528 2 T33527 hypothetical prote  
996 14 53.8 529 2 S46116 probable regulator  
997 14 53.8 530 2 JCS939 estrogen receptor  
998 14 53.8 531 2 T32276 hypothetical prote  
999 14 53.8 532 2 S18448 variant surface g1  
1000 14 53.8 533 2 T31002 hypothetical prote

## ALIGNMENTS

RESULT 1  
S59448  
hypothetical protein YMR206w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: UNIPROT:Q03695; UNIPARC:UP1000013B90A; EMBL:Z48755; NID:G736296; PID  
C:Species: Saccharomyces cerevisiae  
C>Date: 30-Nov-1995 #sequence\_revision 23-Feb-1996 #text\_change 09-Jul-2004  
R:Odell, C.; Bowman, S.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S59441  
A:Accession: S59448  
A:Molecule type: DNA  
A:Residues: 1-313 <ODE>  
A:Cross-references: UNIPROT:Q03695; UNIPARC:UP1000013B90A; EMBL:Z48755; NID:G736296; PID  
A:Experimental source: strain AB972  
C:Genetics:  
A:Gene: MIPS:YMR206w  
A:Cross-References: SGD:S0004819  
A:Map position: 13R  
Query Match 69.2%; Score 18; DB 2; Length 313;  
Best Local Similarity 20.0%; Pred. No. 0.44;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
OY 1 CXXXXXXXXXC 10  
DB 245 CSSSSSSSNC 254  
RESULT 2  
E75433  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zaleski, C.; Ma  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: E75433  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-480 <WHI>  
A:Cross-references: UNIPROT:Q9RV67; UNIPARC:UPI00000C18A8; GB:AE001963; GB:AE000513; NID

A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR1142  
A:Map position: 1  
C:Superfamily: Deinococcus radiodurans hypothetical protein DR1142

Query Match 69.2%; Score 18; DB 2; Length 480;  
Best Local Similarity 20.0%; Pred. No. 0.51;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10  
DB 464 CASAAATTAAC 473

RESULT 3  
J01560  
hypothetical 20.6K protein - Lymantria dispar nuclear polyhedrosis virus  
N:Alternate names: hypothetical protein 4  
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdNPV  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
R:J.Bjornson, R.M.; Rohmann, G.F.  
J. Gen. Virol. 73, 1499-1504, 1992  
A:Title: Nucleotide sequence of the polyhedron envelope protein gene region of the Lyman  
A:Reference number: P00339; MUID:92300345; PMID:1607868  
A:Accession: J01560  
A:Molecule type: DNA  
A:Residues: 1-194 <BJO>  
A:Cross-references: UNIPROT:P36868; UNIPARC:UPI000017A7BF; DBJ:DI0836

Query Match 65.4%; Score 17; DB 2; Length 194;  
Best Local Similarity 20.0%; Pred. No. 1.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10  
DB 74 CSAAMTSSSC 83

RESULT 4  
JC4082  
coat protein - Cymbidium mosaic virus  
C:Species: Cymbidium mosaic virus  
C>Date: 21-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004  
R:Ryu, K.H.; Yoon, K.E.; Park, W.M.  
Gene 156, 303-304, 1995  
A:Title: Cloning and sequencing of a cDNA encoding the coat protein of a Korean isolate  
A:Reference number: JC4082; MUID:95278762; PMID:7758973  
A:Accession: JC4082  
A:Molecule type: mRNA  
A:Residues: 1-220 <RYU>  
A:Cross-references: UNIPROT:Q66152; UNIPARC:UPI00000F22DE; EMBL:X81051; NID:G897718; PID  
A>Note: The authors translated the codon GTG for residue 161 as Leu  
C:Superfamily: potato virus coat protein  
C:Keywords: coat protein

Query Match 65.4%; Score 17; DB 2; Length 220;  
Best Local Similarity 20.0%; Pred. No. 1.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10  
DB 166 CSAATLTATTC 175

RESULT 5  
AB2920  
cobalamin biosynthetic protein [imported] - Agrobacterium tumefaciens (strain C58, Dupon  
C:Species: Agrobacterium tumefaciens  
C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: AB2920

R.Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavain, T.; Levy, R.; Li, M.; McClellan,  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: 100, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AB2920  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-376 <KUR>  
A:Cross-references: UNIPROT:Q8UB06; UNIPARC:UPI0000164772; GB:AE008688; PIDN:AAL43776.1;  
C:Genetics:  
A:Gene: chid  
A:Map position: circular chromosome  
C:Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 65.4%; Score 17; DB 2; Length 376;  
Best Local Similarity 20.0%; Pred. No. 2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
|  
DB 17 CAAATKAC 26

RESULT 6  
cobalamin biosynthetic protein chid (PA2908) [imported] - *Agrobacterium tumefaciens* (str  
C:Species: *Agrobacterium tumefaciens*  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: C97694  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Woliam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: C97694  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-388 <KUR>  
A:Cross-references: UNIPROT:Q8UB06; UNIPARC:UPI00000D2014; GB:AE007869; PIDN:AAK86508.1;  
C:Genetics:  
A:Gene: AGR\_C\_5073  
A:Map position: circular chromosome  
C:Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 65.4%; Score 17; DB 2; Length 388;  
Best Local Similarity 20.0%; Pred. No. 2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
|  
DB 29 CAAATKAC 38

RESULT 7  
T3167  
hypothetical protein KO1C8.2 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T3167  
R:Sim, M.  
submitted to the EMBL Data Library, April 1995  
A:Reference number: Z19702  
A:Accession: T3167  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-389 <WIL>  
A:Cross-references: UNIPROT:Q21081; UNIPARC:UPI0000080155; EMBL:Z49068; PIDN:CAA8685.1;  
A:Experimental source: clone KO1C8

C:Genetics:  
A:Gene: CESP:KO1C8.2  
A:Map position: 2  
A:introns: 54/2; 146/3; 208/3; 283/1; 379/3

Query Match 65.4%; Score 17; DB 2; Length 389;  
Best Local Similarity 20.0%; Pred. No. 2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
|  
DB 123 CSSISTSSC 132

RESULT 8  
T37314  
probable kexin (EC 3.4.21.61) - *Caenorhabditis elegans* (fragment)  
N:Alternate names: Blisterase 4  
C:Species: *Caenorhabditis elegans*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jul-2000  
C:Accession: T37314  
R:Thacker, C.; Peters, K.; Strayko, M.; Rose, A.M.  
Genes Dev. 9, 956-971, 1995  
A:Title: The bil-4 locus of *Caenorhabditis elegans* encodes structurally distinct kex2/su  
A:Reference number: Z21679; MUID:95293228; PMID:7774813  
A:Accession: T37314  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-570 <THA>  
A:Cross-references: UNIPARC:UPI00001688E1; EMBL:L29440; NID:g459702; PIDN:AAA98752.1; PI  
C:Genetics:  
A:Gene: bil-4  
A:Map position: 1  
C:Keywords: alternative splicing; hydrolase; serine proteinase

Query Match 65.4%; Score 17; DB 2; Length 570;  
Best Local Similarity 20.0%; Pred. No. 2.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
|  
DB 311 CTSSSATSC 320

RESULT 9  
S33575  
dnak-type molecular chaperone precursor, mitochondrial - *Leishmania major*  
N:Alternate names: heat shock protein 70-related protein; mitochondrial stress protein  
C:Species: *Leishmania major*  
C:Date: 09-Dec-1997 #sequence\_revision 09-Dec-1997 #text\_change 31-Dec-2004  
C:Accession: S33575; S78090; S05438  
R:Searle, S.; McCrossan, M.V.; Smith, D.F.  
J. Cell Sci. 104, 1091-1100, 1993  
A:Title: Expression of a mitochondrial stress protein in the protozoan parasite *Leishman*  
A:Reference number: S33575; MUID:93300981; PMID:8314893  
A:Accession: S33575  
A:Molecule type: DNA  
A:Residues: 1-634 <SEA>  
A:Cross-references: UNIPROT:P12076; UNIPARC:UPI0000177024; EMBL:X64137  
R:Smith, D.F.  
submitted to the EMBL Data Library, January 1992  
A:Reference number: S78090  
A:Accession: S78090  
A:Molecule type: DNA  
A:Residues: 1-460, 481-500, 'QGERIRASENQIRGEFDSG', 501-634 <SMT>  
A:Cross-references: UNIPARC:UPI000012C5C; EMBL:X64137; NID:g911289; PIDN:CAA45498.1; PI  
R:Searle, S.; Campos, A.J.R.; Coulson, R.M.R.; Spithill, T.W.; Smith, D.F.  
Nucleic Acids Res. 17, 5081-5095, 1989  
A:Title: A family of heat shock protein 70-related genes are expressed in the promastigote  
A:Reference number: S05438; MUID:89345072; PMID:2762121  
A:Accession: S05438  
A:Status: translation not shown  
A:Molecule type: DNA

A:Residues: 1-249 <SEP>  
A:Cross-references: UNIPARC:UPI000016BFLC, EMBL:X14574, NID:g9561, PIDD:CAA32713.1; PIDD:  
C:Genetics:  
A:Gene: lmbp70.1  
A:Genome: nuclear  
C:Function:  
A:Description: involved in protein folding and assembling/disassembling of protein comp  
C:Superfamily: bcr protein  
C:Keywords: ATP; mitochondrion; molecular chaperone; stress-induced protein  
F:1-23/Domain: transit peptide (mitochondrion) #status predicted <TMP>  
F:24-634/Product: heat shock protein 70, mitochondrial #status predicted <MAT>  
  
Query Match 65.4%; Score 17; DB 2; Length 634;  
Best Local Similarity 20.0%; Pred. No. 2.4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXC 10  
DB 7 CGSAAASAC 16  
  
RESULT 10  
D87803  
protein hli-4D [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 21-Jun-2002  
C:Accession: D87803  
C:Anonymous: The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A>Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_elg  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: D87803  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-942 <STO>  
A:Cross-references: UNIPARC:UPI000016B639, GB:chr\_I, PIDD:AAB96754.1, PIDD:g2773243; GSPT  
C:Genetics:  
A:Gene: hli-4D  
A:Map position: 1  
C:Superfamily: kexin; subtilisin homology  
  
Query Match 65.4%; Score 17; DB 2; Length 942;  
Best Local Similarity 20.0%; Pred. No. 2.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXC 10  
DB 683 CTSSSATS 692  
  
RESULT 11  
S09118  
G surface protein 168 - Paramecium primaurelia  
C:Species: Paramecium primaurelia  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: S09118  
R:Prat, A.  
J. Mol. Biol. 211, 521-535, 1990  
A>Title: Conserved sequences flank variable tandem repeats in two alleles of the G surfa  
A:Reference number: S09118; MUID:90172419; PMID:2308165  
A:Accession: S09118  
A:Molecule type: DNA  
A:Residues: 1-2704 <PRA>  
A:Cross-references: UNIPROT:PI7053, UNIPARC:UPI000012AE23, EMBL:X52133, NID:g10049; PIDD:  
C:Genetics:  
A:Genetic code: SGC5  
C:Superfamily: G surface protein  
  
Query Match 65.4%; Score 17; DB 2; Length 2704;  
Best Local Similarity 20.0%; Pred. No. 4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 1632 CVAATPAATTC 1641  
  
RESULT 12  
F84312  
hypothetical protein Vng1598h [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: F84312  
R:Ng, W.V.; Kennedy, S.P.; Mahatras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S  
; Lehnhauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
A>Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: F84312  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-71 <STO>  
A:Cross-references: UNIPROT:Q9HPK1, UNIPARC:UPI0000063920, GB:AE004437, NID:g10581076; P  
C:Genetics:  
A:Gene: VNG1598H  
  
Query Match 61.5%; Score 16; DB 2; Length 71;  
Best Local Similarity 20.0%; Pred. No. 4.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXC 10  
DB 30 CSRSVSTSTC 39  
  
RESULT 13  
E70531  
hypothetical protein Rv2706c - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: E70531  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S  
; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Raglandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: E70531  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-85 <COL>  
A:Cross-references: UNIPROT:O07207, UNIPARC:UPI00000C14DC, GB:Z96072, GB:AL123456; NID:g  
C:Genetics:  
A:Gene: Rv2706c  
  
Query Match 61.5%; Score 16; DB 2; Length 85;  
Best Local Similarity 20.0%; Pred. No. 5;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXC 10  
DB 24 CSATVAAVC 33  
  
RESULT 14  
S51479  
drought-induced protein Di21 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 15-Jul-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S51479; S43176

R.Goelti, F.; Bertauche, N.; Vartanian, N.; Giraudat, J.

Mol. Gen. Genet. 246, 10-18, 1995

A>Title: Abscisic acid-dependent and -independent regulation of gene expression by proga

A.Reference number: S51476; MUID:95124220; PMID:7823504

A.Accession: S51479

A>Status: nucleic acid sequence not shown; translation not shown

A.Molecule type: mRNA

A.Residues: 1-104 <GOS>

A.Cross-references: UNIPROT:Q39084; UNIPARC:UPI000000BFP1; EMBL:X78585; NID:9469111; PID

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994

C.Genetics:

A:Gene: D121

C:Superfamily: late embryogenesis-abundant protein leas

Query Match 61.5%; Score 16; DB 2; Length 104;  
Best Local Similarity 20.0%; Pred. No. 5.4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10  
Db 14 CSAAGSUSC 23

#### RESULT 15

A71249  
hypothetical protein PH0248 - *Pyrococcus horikoshii*

C:Species: *Pyrococcus horikoshii*

C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 12-Jul-2004

C:Accession: A71249

R.Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A>Title: Complete sequence and gene organization of the genome of a hyper-thermophilic e

A.Reference number: A71000; MUID:98344137; PMID:9679194

A.Accession: A71249

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A.Molecule type: DNA

A.Residues: 1-105 <KAW>

A.Cross-references: UNIPROT:Q57986; UNIPARC:UPI0000062D66; GB:AP000001; NID:93236128; PI

A:Experimental source: strain OT3

A>Note: this accession replaces an interim accession for a sequence replaced by GenBank

C.Genetics:

A:Gene: PH0248

Query Match 61.5%; Score 16; DB 2; Length 105;  
Best Local Similarity 20.0%; Pred. No. 5.4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10  
Db 35 CFTASASLAC 44

#### RESULT 16

F72549  
hypothetical protein APE1682 - *Aeropyrum pernix* (strain K1)

C:Species: *Aeropyrum pernix*

C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004

C:Accession: F72549

R.Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Maesuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyr*

A.Reference number: A72450; MUID:99310339; PMID:10382966

A.Accession: F72549

A>Status: preliminary

A.Molecule type: DNA

A.Residues: 1-108 <KAW>

A.Cross-references: UNIPROT:Q9YB85; UNIPARC:UPI000005E048; DDBJ:AP000062; NID:95105244;

A:Experimental source: strain K1

C.Genetics:

A:Gene: APE1682

C:Superfamily: *Aeropyrum pernix* hypothetical protein APE1682

Query Match 61.5%; Score 16; DB 2; Length 108;  
Best Local Similarity 20.0%; Pred. No. 5.5;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10  
Db 8 CFSASKAC 17

#### RESULT 17

A23473

chymotrypsin-like proteinase (EC 3.4.21.-) ~ pig (tentative sequence) (fragments)

N.Alternate names: pancreatic elastase II (identification)

C:Species: *Sus scrofa domestica* (domestic pig)

C>Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 09-Jul-2004

C:Accession: A23473

R.Verred, M.; Gertler, A.; Burstein, Y.

Int. J. Pept. Protein Res. 27, 183-190, 1986

A.Reference number: A23473; MUID:86194934; PMID:3634756

A.Accession: A23473

A.Molecule type: protein

A.Residues: 1-126 <VER>

A:Cross-references: UNIPROT:Q7M325; UNIPARC:UPI0000175C02

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

Query Match 61.5%; Score 16; DB 2; Length 126;  
Best Local Similarity 20.0%; Pred. No. 5.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10  
Db 87 CAGASGASSC 96

#### RESULT 18

T49498

hypothetical protein B14D6.500 [imported] - *Neurospora crassa*

C:Species: *Neurospora crassa*

C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000

C:Accession: T49498

R.Schulte, U.; Algen, V.; Hohlseil, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022

A.Accession: T49498

A>Status: preliminary

A.Molecule type: DNA

A.Residues: 1-129 <SCH>

A:Cross-references: UNIPARC:UPI0000179E1D; EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.500

A:Experimental source: BAC clone B14D6, strain OR74A

C.Genetics:

A:Gene: NCSP:B14D6.500

A:Map position: 6

C:Superfamily: *Neurospora crassa* hypothetical protein B14D6.500

Query Match 61.5%; Score 16; DB 2; Length 129;  
Best Local Similarity 20.0%; Pred. No. 5.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10  
Db 90 CSARASISTC 99

#### RESULT 19

A96746

hypothetical protein T9N14.5 [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: A96746

R.Theologis, A.; Ecker, J.R.; Palm, C.J.; Pedersen, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroe, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 Kerr, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: A96746  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-133 <STO>  
 A:Cross-references: UNIPROT:Q9C7T1, UNIPARC:UPI00000450D; GB:AE005173; NID:G10645374; F  
 C:Genetics:  
 A:Gene: T9N14.5  
 A:Map position: 1

Query Match 61.5%; Score 16; DB 2; Length 133;  
 Best Local Similarity 20.0%; Pred. No. 5.9;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
 DB 93 CYSSTSSVSSC 102

RESULT 20  
 S09762  
 hypothetical protein TRL13 precursor - human cytomegalovirus (strain AD169)  
 C:Species: human cytomegalovirus, human herpesvirus 5  
 A:Note: host Homo sapiens (man)  
 C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
 C:Accession: S09762  
 R:Chen, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Hornselli, T.;  
 M.; Barrell, B.G.  
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990  
 A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus  
 A:Reference number: S09749; MUID:90269039; PMID:2161319  
 A:Accession: S09762  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-147 <CHB>  
 A:Cross-references: UNIPROT:P16811, UNIPARC:UPI000012D86F; EMBL:X17403; NID:G59591; PIDN  
 A:Note: this sequence was submitted to the EMBL Data Library, December 1989  
 C:Superfamily: human cytomegalovirus hypothetical protein TRL13  
 C:Keywords: glycoprotein  
 F:1-35/Domain: signal sequence #status predicted <Sig>  
 F:126-147/Product: hypothetical protein TRL13 #status predicted <MAT>  
 F:21,35,54,62,99,110,125,133,144/Binding site: carbohydrate (Aam) (covalent) #status pre

Query Match 61.5%; Score 16; DB 2; Length 147;  
 Best Local Similarity 20.0%; Pred. No. 6.1;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
 DB 97 CTNTTTTTC 106

RESULT 21  
 PNO103  
 hypothetical 17K protein - barley stripe mosaic virus  
 C:Species: barley stripe mosaic virus, BSMV  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C:Accession: PNO103  
 R:Kozlov, Y.V.; Afanasiev, B.N.; Rupakov, V.V.; Golova, Y.B.; Kulaeva, O.I.; Dolja, V.V.  
 Mol. Biol. (Mosk.) 23, 1080-1090, 1989  
 A:Title: The complete nucleotide sequence of barley stripe mosaic virus RNA 3 and its va  
 A:Reference number: PNO102; MUID:9006400; PMID:2586501  
 A:Accession: PNO103  
 A:Molecule type: genomic RNA  
 A:Residues: 1-153 <KOZ>

A:Cross-references: UNIPROT:Q07118; UNIPARC:UPI00000F58A5  
 C:Genetics:  
 A:Map position: segment 3(III)

Query Match 61.5%; Score 16; DB 2; Length 153;  
 Best Local Similarity 20.0%; Pred. No. 6.2;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
 DB 10 CGTSTSTTC 19

RESULT 22  
 T28088  
 hypothetical protein ZK899.1 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T28088  
 R:Kershaw, J.  
 submitted to the EMBL Data Library, September 1994  
 A:Reference number: Z20468  
 A:Accession: T28088  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-161 <WIL>  
 A:Cross-references: UNIPROT:Q23659, UNIPARC:UPI000007B9F8, EMBL:Z77140; PIDN:CAA85496.1;  
 A:Experimental source: clone ZK899  
 C:Genetics:  
 A:Gene: CESP:ZK899.1  
 A:Map position: X  
 A:Introns: 46/3; 89/3; 144/2  
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein ZK899.1

Query Match 61.5%; Score 16; DB 2; Length 161;  
 Best Local Similarity 20.0%; Pred. No. 6.3;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
 DB 152 C8TESTSTQC 161

RESULT 23  
 JQ1252  
 hypothetical 16.7K protein - chrysanthemum virus B  
 C:Species: chrysanthemum virus B  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C:Accession: JQ1252  
 R:Levay, K.; Zavriv, S.  
 J. Gen. Virol. 72, 2333-2337, 1991  
 A:Title: Nucleotide sequence and gene organization of the 3'-terminal region of chrysant  
 A:Reference number: JQ1246; MUID:92013948; PMID:1919520  
 A:Accession: JQ1252  
 A:Molecule type: genomic RNA  
 A:Residues: 1-164 <LEV>  
 A:Cross-references: UNIPROT:Q7LZW7; UNIPARC:UPI0000179B53

Query Match 61.5%; Score 16; DB 2; Length 164;  
 Best Local Similarity 20.0%; Pred. No. 6.3;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
 DB 80 CSLSSAASC 89

RESULT 24  
 T34277  
 hypothetical protein F46H5.2 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T34277



R.Nhan, M.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid F46H5.  
A:Reference number: Z21498  
A:Accession: T34277  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-229 <RNA>  
A:Cross-references: UNIPROT:Q20490; UNIPARC:UPI0000179CAC; EMBL:U41543; P1DN:AA837021.1;  
C:Genetics:  
A:Gene: CESP:F46H5.2  
A:Map position: X  
A:introns: 16/1; 85/2; 151/2

Query Match 61.5%; Score 16; DB 2; Length 229;  
Best Local Similarity 20.0%; Pred. No. 7.1;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 23 CEASASNSSC 32

RESULT 25  
A45724  
pectate lyase (EC 4.2.2.2) - fungus (Fusarium solani)  
C:Species: Fusarium solani  
C>Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A45724  
R:Gonzalez-Candelas, L.; Kolatukudy, P.E.  
J. Bacteriol. 174, 6343-6349, 1992  
A:Title: Isolation and analysis of a novel inducible pectate lyase gene from the phytos  
A:Reference number: A45724; MUID:93015682; PMID:1400187  
A:Accession: A45724  
A:Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-242 <CON>  
A:Cross-references: UNIPROT:Q04701; UNIPARC:UPI000006C8B5; GB:M94691; NID:g168155; P1DN:  
A:Experimental source: isolate T8  
A:Note: sequence extracted from NCBI backbone (NCBIN:115473, NCBIPI:115474)  
C:Keywords: carbon-oxygen lyase

Query Match 61.5%; Score 16; DB 2; Length 242;  
Best Local Similarity 20.0%; Pred. No. 7.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 233 CTVGTSTTSC 242

RESULT 26  
S72219  
chymotrypsin B - Atlantic cod (Fragments)  
C:Species: Gadus morhua (Atlantic cod)  
C>Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 09-Jul-2004  
C:Accession: S72219  
R:Leeth-Larsen, R.; Asgeirsson, B.; Thorolfsson, M.; Norregard-Madsen, M.; Hojrup, P.  
Biochim. Biophys. Acta 1297, 49-56, 1996  
A:Title: Structure of chymotrypsin variant B from Atlantic cod, Gadus morhua.  
A:Reference number: S72219; MUID:96439045; PMID:8841380  
A:Accession: S72219  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-14,15-24 <DET>  
A:Cross-references: UNIPROT:Q9PM06; UNIPARC:UPI0000175C00; UNIPARC:UPI0000175C01  
C:Superfamily: trypsin; trypsin homology  
F:15-237/Domain: trypsin homology <TRY>

Query Match 61.5%; Score 16; DB 2; Length 244;  
Best Local Similarity 20.0%; Pred. No. 7.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 181 CAGAGATSC 190

RESULT 27  
A55035  
cysteine-rich protein CRP1 - earthworm (Eudryotreus buchholzi)  
C:Species: Eudryotreus buchholzi  
C>Date: 14-Nov-1994 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: A55035; S45034  
R:Willuhn, J.; Schmitt-Wrede, H.P.; Greven, H.; Wunderlich, F.  
J. Biol. Chem. 269, 24688-24691, 1994  
A:Title: cDNA cloning of a cadmium-inducible mRNA encoding a novel cysteine-rich, non-m  
A:Reference number: A55035; MUID:95014230; PMID:7929141  
A:Accession: A55035  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-251 <WIL>  
C:Superfamily: ultra-high-sulfur keratin

Query Match 61.5%; Score 16; DB 2; Length 251;  
Best Local Similarity 20.0%; Pred. No. 7.4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 28 CGTSTTAGSC 37

RESULT 28  
J01724  
EI membrane glycoprotein precursor - canine coronavirus (strain Inavc-1)  
N:Alternate names: matrix glycoprotein  
C:Species: canine coronavirus  
C>Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 09-Jul-2004  
C:Accession: J01724  
R:Horsburgh, B.C.; Briarley, I.; Brown, T.D.K.  
J. Gen. Virol. 73, 2849-2862, 1992  
A:Title: Analysis of a 9.6 kb sequence from the 3' end of canine coronavirus genomic RN  
A:Reference number: P00481; MUID:93057357; PMID:1431811  
A:Accession: J01724  
A:Molecule type: genomic RNA  
A:Residues: 1-262 <HOR>  
A:Cross-references: UNIPROT:P36299; UNIPARC:UPI0000138986; DDBJ:DI3096; NID:g406193; P1  
C:Genetics:  
A:Gene: M  
C:Superfamily: coronavirus EI membrane glycoprotein  
C:Keywords: glycoprotein; matrix protein; transmembrane protein  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-262/Domain: signal sequence #status predicted <SIG>  
F:56-72/Domain: transmembrane #status predicted <TM1>  
F:85-101/Domain: transmembrane #status predicted <TM2>  
F:115-134/Domain: transmembrane #status predicted <TM3>  
F:32-55/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:32-57/Binding site: carbohydrate (Ser) (covalent) #status predicted  
F:44/Binding site: carbohydrate (Thr) (covalent) #status predicted

Query Match 61.5%; Score 16; DB 1; Length 262;  
Best Local Similarity 20.0%; Pred. No. 7.5;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 21 CAMTESSTSC 30

RESULT 29  
S47537  
chymotrypsin (EC 3.4.21.1) precursor - Atlantic cod  
C:Species: Gadus morhua (Atlantic cod)

C.Date: 26-Dec-1994 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004  
C.Accession: S47537; S43163  
R.Gutmundsdottir, A.; Oskarsson, S.; Eakin, A.E.; Craik, C.S.; Bjarnason, J.B.  
Biochim. Biophys. Acta 1219, 211-214, 1994  
A.Title: Atlantic cod cDNA encoding a psychrophilic chymotrypsinogen.  
A.Reference number: S47537; MUID:94368660; PMID:8086467  
A.Accession: S47537  
A.Molecule type: mRNA  
A.Residues: 1-263 <GUD>  
A.Cross-references: UNIPROT:P47796; UNIPARC:UPI0000128664; EMBL:X78490; NID:G468750; PID  
C.Superfamily: trypsin; trypsin homology  
C.Keywords: hydrolase; protein digestion; serine proteinase  
P.1-18/Domain: signal sequence #status predicted <SIG>  
P.19-263/Product: chymotrypsin #status predicted <MAT>  
P.34-256/Domain: trypsin homology <TRY>  
P.75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 61.5%; Score 16; DB 2; Length 263;  
Best Local Similarity 20.0%; Pred. No. 7.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10  
DB 200 CAGAGAGASC 209

RESULT 30  
A85856  
probable elongation factor yelp [imported] - Escherichia coli (strain O157:H7, substrain  
C.Species: Escherichia coli  
C.Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 05-Oct-2004  
C.Accession: A85856  
R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Iller, L.; Grobbeck, E.J.; Davis, N.W.; Llin, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Natter, A.; Grotbeck, E.J.; Davis, N.W.; Llin, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
A.Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A.Reference number: A85480; MUID:21074935; PMID:11206551  
A.Accession: A85856  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-275 <STO>  
A.Cross-references: UNIPARC:UPI00000020288; GB:AE005174; NID:G12516497; PIDN:AA657309.1;  
A.Experimental source: strain O157:H7, substrain EDP933  
C.Genetics:  
A.Gene: yelp  
C.Superfamily: translation elongation factor P

Query Match 61.5%; Score 16; DB 2; Length 275;  
Best Local Similarity 20.0%; Pred. No. 7.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10  
DB 54 CTSKSANTTC 63

RESULT 31  
B64986  
hypothetical 30.9 kD protein in frub 5' region - Escherichia coli (strain K-12)  
C.Species: Escherichia coli  
C.Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 05-Oct-2004  
C.Accession: B64986  
R.Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A.Title: The complete genome sequence of Escherichia coli K-12.  
A.Reference number: A64720; MUID:97426617; PMID:9278503  
A.Accession: B64986  
A.Status: preliminary; nucleic acid sequence not shown; translation not shown  
A.Molecule type: DNA  
A.Residues: 1-275 <BLAT>  
A.Cross-references: UNIPARC:UPI0000167DCL; GB:AE000306; GB:U00096; NID:G1788489; PIDN:AA  
A.Experimental source: strain K-12, substrain MG1655

C.Genetics:  
A.Gene: yelp  
C.Superfamily: translation elongation factor P

Query Match 61.5%; Score 16; DB 2; Length 275;  
Best Local Similarity 20.0%; Pred. No. 7.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10  
DB 54 CTSKSANTTC 63

RESULT 32  
G91011  
probable elongation factor [imported] - Escherichia coli (strain O157:H7, substrain RMD  
C.Species: Escherichia coli  
C.Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 05-Oct-2004  
C.Accession: G91011  
R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A.Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
A.Reference number: A99629; MUID:21156231; PMID:11258796  
A.Accession: G91011  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-275 <HAY>  
A.Cross-references: UNIPARC:UPI00000020288; GB:BA000007; PIDN:BA836486.1; PID:G13362532;  
A.Experimental source: strain O157:H7, substrain RMD 0509952  
C.Genetics:  
A.Gene: EC93063  
C.Superfamily: translation elongation factor P

Query Match 61.5%; Score 16; DB 2; Length 275;  
Best Local Similarity 20.0%; Pred. No. 7.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10  
DB 54 CTSKSANTTC 63

RESULT 33  
T23682  
hypothetical protein M02G9.3 - Caenorhabditis elegans  
C.Species: Caenorhabditis elegans  
C.Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C.Accession: T23682  
R.Matthews, L.  
submitted to the EMBL Data Library, November 1996  
A.Reference number: Z19781  
A.Accession: T23682  
A.Status: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: DNA  
A.Residues: 1-294 <WIL>  
A.Cross-references: UNIPROT:O9XUS0; UNIPARC:UPI0000076A8F; EMBL:Z81573; PIDN:CAB04626.1;  
A.Experimental source: clone M02G9  
C.Genetics:  
A.Gene: CESP:M02G9.3  
A.Map position: 2  
A.Introns: 20/3; 76/2; 182/3; 223/3

Query Match 61.5%; Score 16; DB 2; Length 294;  
Best Local Similarity 20.0%; Pred. No. 7.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 10  
DB 43 CSSSSNSNYC 52

RESULT 34

C71498  
probable yop translocation R - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C/Species: Chlamydia trachomatis  
C/Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jul-2004  
C/Accession: C71498  
R/Stephen, R.S.; Kallman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
A/Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
A/Reference number: A71570; MWID:9900809; PMID:9784136  
A/Accession: C71498  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-306 <RN>  
A/Cross-references: UNIPROT:084566; UNIPARC:UPI00000D3360; GB:AE001327; GB:AE001273; NID: A/Experimental source: serotype D, strain UW-3/Cx  
C/Genetics:  
A/Genes: yscR

Query Match 61.5%; Score 16; DB 2; Length 306;  
Best Local Similarity 20.0%; Pred. No. 7.9;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10  
DB 24 CSNAGASSC 33

## RESULT 35

T00500  
probable elicitor response element-binding protein WRKY3 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 31-Dec-2004  
C/Accession: T00500; C84623  
R/Rounsley, S.D.; Lin, X.; Kechum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul submitted to the EMBL Data Library, November 1997  
A/Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.  
A/Reference number: Z14164  
A/Accession: T00500  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-317 <RN>  
A/Cross-references: UNIPROT:022176; UNIPARC:UPI0000138FTD; EMBL:AC002391; NID:G2642427;  
A/Experimental source: cultivar Columbia  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MWID:20083487; PMID:10617197  
A/Accession: C84623  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-317 <STO>  
A/Cross-references: UNIPARC:UPI0000138FTD; GB:AE002093; NID:G2642432; PIDN:AAB87100.1; C/Genetics:  
A/Genes: WRKY3; ATSP:T20D16.5; At2g23320  
A/Map position: 2  
A/Introns: 218/2; 260/2  
C/Superfamily: DNA-binding protein WRKY3

Query Match 61.5%; Score 16; DB 2; Length 317;  
Best Local Similarity 20.0%; Pred. No. 8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10  
DB 203 CASASSGRC 122

## RESULT 36

A84792  
hypothetical protein At2g37380 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: A84792  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MWID:20083487; PMID:10617197  
A/Accession: A84792  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-321 <STO>  
A/Cross-references: UNIPROT:09ZUS8; UNIPARC:UPI00000A2519; GB:AE002093; NID:G4056491; C/Genetics:  
A/Genes: At2g37380  
A/Map position: 2  
C/Superfamily: Arabidopsis thaliana hypothetical protein At2g39370

Query Match 61.5%; Score 16; DB 2; Length 321;  
Best Local Similarity 20.0%; Pred. No. 8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10  
DB 306 CSRTSVSTC 315

## RESULT 37

J08202  
paired-box-containing Pax4 protein - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 09-Jul-2004  
C/Accession: J08202  
R/Tokuyama, Y.; Yagui, K.; Sakurai, K.; Hashimoto, N.; Saico, Y.; Kanatsuka, A. Biochem. Biophys. Res. Commun. 248, 153-156, 1998  
A/Title: Molecular cloning of rat Pax4: Identification of four isoforms in rat insulinoma  
A/Reference number: J08202; MWID:98340866; PMID:9675102  
A/Accession: J08202  
A/Molecule type: mRNA  
A/Residues: 1-349 <TOX>  
A/Cross-references: UNIPROT:088436; UNIPARC:UPI00001136B; GB:AF053100; NID:G3192973; P C/Superfamily: paired box transcription factor Pax-4; homeobox homology; paired box hom C/Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F/5-129/Domain: paired box homology <PBH>  
F/171-227/Domain: homeobox homology <HOX>

Query Match 61.5%; Score 16; DB 2; Length 349;  
Best Local Similarity 20.0%; Pred. No. 8.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10  
DB 290 CSSTTSQAC 299

## RESULT 38

T04007  
hypothetical protein T5L19.160 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 09-Jul-2004  
C/Accession: T04007  
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X submitted to the Protein Sequence Database, March 1999  
A/Reference number: Z15184  
A/Accession: T04007  
A/Molecule type: DNA  
A/Residues: 1-380 <BEV>  
A/Cross-references: UNIPROT:09T0G1; UNIPARC:UPI00000A0088; EMBL:AL049481  
A/Experimental source: cultivar Columbia; BAC clone T5L19  
C/Genetics:  
A/Map position: 4  
A/Introns: 38/3; 83/3; 92/3; 123/1; 137/3; 170/3; 190/1; 207/3; 241/3; 263/3; 301/2; 358 A/Note: T5L19.160

```
Query Match          61.5%; Score 16; DB 2; Length 380;
Best Local Similarity 20.0%; Pred. No. 8.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 6 CSTASSSGGC 15

RESULT 39
T26077
hypohectical protein W02A2.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26077
R:Ainscough, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z20148
A:Accession: T26077
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-397 <MIL>
A:Cross-references: UNIPROT:Q9XUB6; UNIPARC:UPI0000076467; EMBL:Z82286; PIDs: CAB05306.1;
A:Experimental source: clone W02A2
C:Genetics:
A:Gene: CESP:W02A2.3
A:Map position: 4
A:introns: 15/1; 67/2; 101/3; 259/1; 383/3
C:Superfamily: gliadin

Query Match          61.5%; Score 16; DB 2; Length 397;
Best Local Similarity 20.0%; Pred. No. 8.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 10 CATTAAASC 19

RESULT 40
S41514
RAD52 protein homolog - chicken
C:Species: Gallus gallus (chicken)
C:Date: 25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 31-Dec-2004
C:Accession: S41514
R:Bezubova, O.Y.; Schmidt, H.; Ostermann, K.; Heyer, W.D.; Buerstedde, J.M.
Nucleic Acids Res. 21, 5945-5949, 1993
A>Title: Identification of a chicken RAD52 homologue suggests conservation of the RAD52
A:Reference number: S41514; MUID:94119695; PMID:8290357
A:Accession: S41514
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-422 <BEZ>
A:Cross-references: UNIPROT:P39022; UNIPARC:UPI0000133024
C:Superfamily: human nuclear cap-binding protein

Query Match          61.5%; Score 16; DB 2; Length 422;
Best Local Similarity 20.0%; Pred. No. 8.8;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 17 CTSTNSVAC 26

RESULT 41
C81039
lipopolysaccharide biosynthesis protein, probable NMB1818 [imported] - Neisseria meningi
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: C81039
R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
```

```
Hikey, B.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamshayan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: C81039
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <TEF>
A:Cross-references: UNIPROT:Q9UXZ7; UNIPARC:UPI000030867; GB:AE002531; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1818

Query Match          61.5%; Score 16; DB 2; Length 473;
Best Local Similarity 20.0%; Pred. No. 9.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 421 CTTSSAATC 430

RESULT 42
C81984
probable lipopolysaccharide biosynthesis translocase NMA0643 [imported] - Neisseria meni
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: C81984
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: C81984
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <PAR>
A:Cross-references: UNIPROT:Q9UW9; UNIPARC:UPI0000030868; GB:AL162753; GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0641; NMA0643

Query Match          61.5%; Score 16; DB 2; Length 473;
Best Local Similarity 20.0%; Pred. No. 9.2;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10
Db 421 CLASSAATC 430

RESULT 43
T36342
probable glutamate decarboxylase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36342
R:Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z21575
A:Accession: T36342
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-475 <OLI>
A:Cross-references: UNIPROT:Q9X8J5; UNIPARC:UPI00000808A; EMBL:AL049841; PIDs: CAB42769.
A:Experimental source: strain A3 (2)
C:Genetics:
A:Gene: gad; SCOE:DB:SCB9.23
C:Superfamily: Escherichia coli glutamate decarboxylase

Query Match          61.5%; Score 16; DB 2; Length 475;
```

Best Local Similarity 20.0%; Pred. No. 9.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 131 CSTTGSSEAC 140

RESULT 44

B86460 hypothetical protein F14M2.19 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: B86460

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maitl, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,

Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; PMID:21016719; PMID:11130712

A:Accession: B86460

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-486 <STO>

A:Cross-references: UNIPROT:Q9LQ21; UNIPARC:UPI00000A7F31; GB:AE005172; NID:G9665100; FI

A:Map position: 1

Query Match 61.5%; Score 16; DB 2; Length 486;  
Best Local Similarity 20.0%; Pred. No. 9.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 48 CUSSTGAATC 57

RESULT 45

A55207

glycerol-3-phosphate dehydrogenase gldp (EC 1.1.-.-) - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa

C:Date: 05-May-1995 #sequence\_revision 05-May-1995 #text\_change 09-Jul-2004

C:Accession: A55207

R:Schweizer, H.P.; Po, C.

J. Bacteriol. 176, 2184-2193, 1994

A:Title: Cloning and nucleotide sequence of the gldp gene encoding sn-glycerol-3-phosphat

A:Reference number: A55207; PMID:94209216; PMID:8157588

A:Accession: A55207

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-510 <SCH>

A:Cross-references: UNIPROT:P52111; UNIPARC:UPI000016FCAD; GB:L06231; NID:G450377; PIND:

C:Genetics:

A:Gene: gldp

C:Superfamily: glycerol-3-phosphate dehydrogenase (aerobic)

C:Keywords: oxidoreductase

Query Match 61.5%; Score 16; DB 2; Length 510;  
Best Local Similarity 20.0%; Pred. No. 9.4;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 194 CTMSATAAC 203

RESULT 46

G31277

guinate transport protein - Neurospora crassa (tentative sequence)

N:Alternate names: guinate transporter

C:Species: Neurospora crassa

C:Date: 26-Apr-1989 #sequence\_revision 26-Apr-1989 #text\_change 09-Jul-2004

C:Accession: S04254; G31277

R:Geever, R.F.; Hulet, L.; Baum, J.A.; Tyler, B.M.; Patel, V.B.; Rutledge, B.J.; Case, J.

J. Mol. Biol. 207, 15-34, 1989

A:Title: DNA sequence, organization and regulation of the ga gene cluster of Neurospora

A:Reference number: S04250; PMID:89293848; PMID:2525625

A:Accession: S04254

A:Molecule type: DNA

A:Residues: 1-537 <GE2>

A:Cross-references: UNIPROT:P11636; UNIPARC:UPI0000132E4B; EMBL:X14603; NID:G3060; PIND

C:Genetics:

A:Gene: ga-Y

C:Superfamily: maltose transport protein MAL61

C:Keywords: transmembrane protein

F:22-42/Domain: transmembrane #status predicted <TM01>

F:57-87/Domain: transmembrane #status predicted <TM02>

F:99-119/Domain: transmembrane #status predicted <TM03>

F:132-152/Domain: transmembrane #status predicted <TM04>

F:161-181/Domain: transmembrane #status predicted <TM05>

F:195-215/Domain: transmembrane #status predicted <TM06>

F:286-306/Domain: transmembrane #status predicted <TM07>

F:324-344/Domain: transmembrane #status predicted <TM08>

F:356-376/Domain: transmembrane #status predicted <TM09>

F:390-410/Domain: transmembrane #status predicted <TM10>

F:427-447/Domain: transmembrane #status predicted <TM11>

F:459-479/Domain: transmembrane #status predicted <TM12>

Query Match 61.5%; Score 16; DB 2; Length 537;  
Best Local Similarity 20.0%; Pred. No. 9.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 24 CAAIASFASC 33

RESULT 47

B38418 jockey protein 1 - fruit fly (Drosophila funebris)

C:Species: Drosophila funebris

C:Date: 28-Jun-1991 #sequence\_revision 28-Jun-1991 #text\_change 21-Feb-1997

C:Accession: B38418

R:Wizokhi, L.U.; Mazo, A.M.

Proc. Natl. Acad. Sci. U.S.A. 87, 9216-9220, 1990

A:Title: Evidence for horizontal transmission of the mobile element jockey between diapa

A:Reference number: A38418; PMID:91067678; PMID:1701254

A:Accession: B38418

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-580 <ML2>

A:Cross-references: UNIPARC:UPI000017BE6C; GB:M38437

C:Genetics:

A:Gene: FlyBase:Dfunt/jockey

A:Cross-references: FlyBase:FBgn0012310

Query Match 61.5%; Score 16; DB 2; Length 580;  
Best Local Similarity 20.0%; Pred. No. 9.9;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10

DB 78 CTVTTSQASC 87

RESULT 48

T22518

hypothetical protein F52H3.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T22518

R.Gardner, A.  
 Submitted to the EMBL Data Library, October 1995  
 A:Reference number: Z19575  
 A:Accession: T22518  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-638 <WIL>  
 A:Cross-references: UNIPROT:Q20680; UNIPARC:UPI000013C10D; EMBL:Z66512; PIDD:CAA91322.1;  
 A:Experimental source: clone F52H3  
 C:Genetics:  
 A:Gene: CESP:F52H3.2  
 A:Map position: 2  
 A:Introns: 67/1; 157/2; 285/2; 420/1; 566/3  
 C:Superfamily: gida protein

Query Match 61.5%; Score 16; DB 2; Length 638;  
 Best Local Similarity 20.0%; Pred. No. 10;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXXC 10  
 Db 29 CESAAAAARC 38

RESULT 49  
 C81345  
 Hypothetical protein Cj0736 [imported] - Campylobacter jejuni (strain NCTC 11168)  
 C:Species: Campylobacter jejuni  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
 C:Accession: C81345  
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kellay, J.M.; Churcher, C.; Basham, D.; Chilling  
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell  
 Nature 403, 665-668, 2000  
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
 A:Reference number: A81250; MUID:20150912; PMID:10688204  
 A:Accession: C81345  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-642 <PAR>  
 A:Cross-references: UNIPROT:Q9PPH1; UNIPARC:UPI00000C1D12; GB:AL139076; GB:AL111168; NID  
 A:Experimental source: serotype O2, strain NCTC 11168  
 C:Genetics:  
 A:Gene: Cj0736

Query Match 61.5%; Score 16; DB 2; Length 642;  
 Best Local Similarity 20.0%; Pred. No. 10;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXXC 10  
 Db 375 CSSAAVSSC 384

RESULT 50  
 T37581  
 probable serine-threonine-protein kinase - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T37581  
 R:Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, April 1996  
 A:Reference number: Z21726  
 A:Accession: T37581  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-648 <BAD>  
 A:Cross-references: UNIPARC:UPI000017B219; EMBL:Z70721; PIDD:CAA94704.1; GSPDB:GN00066;  
 C:Genetics:  
 A:Gene: SPDB:SPAC12B10.14C  
 A:Map position: 1

Query Match 61.5%; Score 16; DB 2; Length 648;

Best Local Similarity 20.0%; Pred. No. 10;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXXC 10  
 Db 21 COSNASSSAC 30

Search completed: January 4, 2006, 16:10:29  
 Job time : 12.3478 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 65.3043 Seconds  
(without alignments)  
108.037 Million cell updates/sec

Title: US-09-932-322-11

Perfect score: 26

Sequence: 1 CXXXXXXXXX 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : UniProt\_05.80:\*  
1: uniprot\_prot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	69.2	213	2 Q96KM3 HUMAN	Q96KM3 homo sapien
2	18	69.2	263	2 Q58F4 YTRYP	Q58F4 trypanosoma
3	18	69.2	269	2 Q7PRO7 ANOGA	Q7PRO7 anopheles g
4	18	69.2	281	2 Q5TX10 ANOGA	Q5TX10 anopheles g
5	18	69.2	313	1 YME8 YEAST	YME8 saccharomyc
6	18	69.2	357	2 Q6ZDR4 ORYSA	Q6ZDR4 oryza sativ
7	18	69.2	362	2 Q7QJL0 ANOGA	Q7QJL0 anopheles g
8	18	69.2	480	2 Q9RV87 DEIRA	Q9RV87 deinococcus
9	18	69.2	653	1 TMCC1 HUMAN	TMCC1 homo sapien
10	18	69.2	1005	2 Q5H317 XANOR	Q5H317 xanthomonas
11	18	69.2	1061	2 Q4QCK7 LEIMA	Q4QCK7 leishmania
12	17	65.4	52	2 Q8LPD6 HORVU	Q8LPD6 hordeum vul
13	17	65.4	122	2 Q5DCN3 SCHJA	Q5DCN3 schistosoma
14	17	65.4	128	2 Q96LJ4 HUMAN	Q96LJ4 homo sapien
15	17	65.4	139	2 Q6ZKM7 ORYSA	Q6ZKM7 oryza sativ
16	17	65.4	155	2 Q6H482 ORYSA	Q6H482 oryza sativ
17	17	65.4	160	2 Q657M0 ORYSA	Q657M0 oryza sativ
18	17	65.4	160	2 Q4T4H1 TETNG	Q4T4H1 tetraodon n
19	17	65.4	219	2 Q825E8 STEPM	Q825E8 streptomyce
20	17	65.4	220	2 Q66152 YVIRU	Q66152 cybidium m
21	17	65.4	256	2 Q6YTS2 ORYSA	Q6YTS2 oryza sativ
22	17	65.4	260	2 Q850Z3 ORYSA	Q850Z3 oryza sativ
23	17	65.4	304	2 Q5TVE4 ANOGA	Q5TVE4 anopheles g
24	17	65.4	344	2 Q582B8 YTRYP	Q582B8 trypanosoma
25	17	65.4	361	2 Q57220 YTRYP	Q57220 trypanosoma
26	17	65.4	364	2 Q4YW87 PLABP	Q4YW87 plasmodium
27	17	65.4	368	2 Q70RD3 GEBHY	Q70RD3 gerbera hyb
28	17	65.4	376	1 CBID AGRTS	CBID agrobacteri
29	17	65.4	388	2 Q7PY13 ANOGA	Q7PY13 anopheles g
30	17	65.4	389	2 Q21081 CAEBL	Q21081 caenorhabdi
31	17	65.4	411	2 Q67U23 ORYSA	Q67U23 oryza sativ

32	17	65.4	437	2 Q414Q8 GIBBE	Q414Q8 gibberella
33	17	65.4	471	2 Q9VMG7 DROME	Q9VMG7 drosophila
34	17	65.4	475	2 Q7X7A4 ORYSA	Q7X7A4 oryza sativ
35	17	65.4	491	2 Q4P4C7 USTMA	Q4P4C7 ustilago ma
36	17	65.4	520	2 Q4XZV5 PLACH	Q4XZV5 plasmodium
37	17	65.4	556	2 Q803D5 BRABE	Q803D5 brachydanio
38	17	65.4	568	2 Q5GXM1 XANOR	Q5GXM1 xanthomonas
39	17	65.4	626	2 Q4N2Y7 THEPA	Q4N2Y7 thellesia p
40	17	65.4	630	2 Q69TY6 ORYSA	Q69TY6 oryza sativ
41	17	65.4	634	1 HSP71 LEIMA	HSP71 leishmania
42	17	65.4	635	2 Q4Q747 LEIMA	Q4Q747 leishmania
43	17	65.4	652	2 Q4Q740 LEIMA	Q4Q740 leishmania
44	17	65.4	660	2 Q4Q744 LEIMA	Q4Q744 leishmania
45	17	65.4	662	2 Q4Q745 LEIMA	Q4Q745 leishmania
46	17	65.4	691	2 Q60E18 ORYSA	Q60E18 oryza sativ
47	17	65.4	695	2 Q819P8 YTRYP	Q819P8 trypanosoma
48	17	65.4	709	2 Q516X2 CHAB	Q516X2 chlamydomon
49	17	65.4	734	2 Q9NKE3 DROME	Q9NKE3 drosophila
50	17	65.4	784	2 Q4YNR2 PLABE	Q4YNR2 plasmodium
51	17	65.4	843	2 Q80GV2 HEPVO	Q80GV2 hepatitis b
52	17	65.4	861	2 Q55NLS CRYNE	Q55NLS cryptococcu
53	17	65.4	943	1 BLI4 CAEBL	BLI4 caenorhabdi
54	17	65.4	1089	2 Q4Q9C2 LEIMA	Q4Q9C2 leishmania
55	17	65.4	1103	2 Q4PDJ1 USTMA	Q4PDJ1 ustilago ma
56	17	65.4	1200	2 Q4RF81 TETNG	Q4RF81 tetraodon n
57	17	65.4	1347	2 Q6CC16 YARLI	Q6CC16 yarrowia li
58	17	65.4	1749	2 Q4V217 BURMA	Q4V217 burkholderi
59	17	65.4	1835	2 Q4S914 TETNG	Q4S914 tetraodon n
60	17	65.4	2704	1 G168 PARP	G168 paramecium
61	17	65.4	2717	1 Q94710 PARTE	Q94710 paramecium
62	17	65.4	2721	2 Q6P973 PARTE	Q6P973 paramecium
63	17	65.4	2729	2 Q6P0K6 PARTE	Q6P0K6 paramecium
64	16	61.5	68	2 Q6JL92 NEIGO	Q6JL92 neisseria g
65	16	61.5	71	2 Q9HPK1 HALBA	Q9HPK1 halobacteri
66	16	61.5	80	2 Q6IGD0 DROME	Q6IGD0 drosophila
67	16	61.5	85	2 Q07207 MYCBO	Q07207 mycobacteri
68	16	61.5	85	2 Q7YV24 MYCBO	Q7YV24 mycobacteri
69	16	61.5	85	2 Q7NSB4 CHRVO	Q7NSB4 chromobacte
70	16	61.5	88	2 Q5Z713 ORYSA	Q5Z713 oryza sativ
71	16	61.5	98	2 Q5Z2E5 NOCPA	Q5Z2E5 nocardia fa
72	16	61.5	103	2 Q5YF63 YVIRU	Q5YF63 rock bream
73	16	61.5	104	2 Q39084 ARATH	Q39084 arabidopsis
74	16	61.5	105	2 Q5Y986 PYRHO	Q5Y986 pyrobacateri
75	16	61.5	108	2 Q5Y986 ASRBE	Q5Y986 aspergillus p
76	16	61.5	108	2 Q854Z9 PCALD	Q854Z9 pyriculari
77	16	61.5	108	2 Q6Y9W8 LIGOV	Q6Y9W8 ligustrum o
78	16	61.5	110	2 Q7PJV4 ANOGA	Q7PJV4 anopheles g
79	16	61.5	112	2 Q95M57 BOVIN	Q95M57 bos taurus
80	16	61.5	112	2 Q8V6M9 YVIRU	Q8V6M9 haovirus h
81	16	61.5	120	2 Q6ID11 ARATH	Q6ID11 arabidopsis
82	16	61.5	123	2 Q84WT9 ARATH	Q84WT9 arabidopsis
83	16	61.5	125	2 Q59Z29 CANAL	Q59Z29 candida alb
84	16	61.5	126	2 Q7M325 PIG	Q7M325 sus scrofa
85	16	61.5	126	2 Q6YUC3 ORYSA	Q6YUC3 oryza sativ
86	16	61.5	126	2 Q8CLV7 YERPE	Q8CLV7 yersinia pe
87	16	61.5	129	2 Q7RUV2 NEUTR	Q7RUV2 neurospora
88	16	61.5	133	2 Q6UNK8 BOMMO	Q6UNK8 bombyx mori
89	16	61.5	133	2 Q9C7T1 ARATH	Q9C7T1 arabidopsis
90	16	61.5	137	2 Q9HAI1 HUMAN	Q9HAI1 homo sapien
91	16	61.5	137	2 Q24987 GIALA	Q24987 gladiola lam
92	16	61.5	137	2 Q8MLI9 DROME	Q8MLI9 drosophila
93	16	61.5	142	2 Q08497 HUMAN	Q08497 homo sapien
94	16	61.5	147	1 TR13 HGWYA	TR13 human cytom
95	16	61.5	147	2 Q6RE78 GNOC	Q6RE78 rhodococcus
96	16	61.5	149	2 Q4RZ65 TETNG	Q4RZ65 tetraodon n
97	16	61.5	149	2 Q4SZ33 TETNG	Q4SZ33 tetraodon n
98	16	61.5	151	2 Q5H3T3 XANOR	Q5H3T3 xanthomonas
99	16	61.5	153	1 GAMB BSNV	GAMB brachydanio
100	16	61.5	157	2 Q8LI53 ORYSA	Q8LI53 oryza sativ
101	16	61.5	158	2 Q8CCG9 MOUSE	Q8CCG9 mus musculu
102	16	61.5	161	2 Q23659 CAEBL	Q23659 caenorhabdi
103	16	61.5	162	2 Q91G76 ADEP3	Q91G76 porcine ade
104	16	61.5	164	2 Q7LZW7 CVB	Q7LZW7 chrysaenthem

105	16	61.5	166	2	061045_DROME	061145_drosophila	178	16	61.5	272	2	0655G8_9CNIID	0655g8 nematostell
106	16	61.5	167	2	024969_GIALA	024969 giardia lam	179	16	61.5	272	2	06W6J9_9CNIID	06w6j9 nematostell
107	16	61.5	170	2	09B1L7_GIALA	09b1l7 giardia lam	180	16	61.5	276	2	04W1J3_ASPFU	04w1j3 aspergillus
108	16	61.5	170	2	08X8Z7_RALSTONIA	08x8z7 ralstonia s	181	16	61.5	280	2	06ZRU6_HUMAN	06zr16 homo sapien
109	16	61.5	171	2	06ZP60_HUMAN	06z60 homo sapien	182	16	61.5	280	2	06ZUB8_HUMAN	06z8 homo sapien
110	16	61.5	171	2	060ZL5_CABER	060z5 caenorhabdi	183	16	61.5	281	2	07ZHQ7_9H1V1	07zhq7 human immun
111	16	61.5	171	2	08W0A0_CABEL	08w0a0 caenorhabdi	184	16	61.5	282	1	KR101_HUMAN	P60331 homo sapien
112	16	61.5	171	2	058S81_PSEFU	058s81 pseudomonas	185	16	61.5	284	2	08MIV8_DROME	08m18 drosophila
113	16	61.5	172	2	098EY7_RHILLO	098ey7 rhizobium l	186	16	61.5	285	2	08S902_SOYBN	08s902 glycine max
114	16	61.5	174	2	071D98_DROYA	071d98 drosophila	187	16	61.5	287	2	08WVJ7_PASCI	08wv7 botientia vi
115	16	61.5	175	2	09VS07_DROME	09v07 drosophila	188	16	61.5	287	2	04X7N6_PLACH	04x7n6 plasmodium
116	16	61.5	175	2	04JMM7_9BACT	04jmm7 uncultured	189	16	61.5	287	2	05U140_DROME	05u140 drosophila
117	16	61.5	179	2	04M046_9BURK	04m046 burkholderi	190	16	61.5	290	2	08EP14_DROME	08ep14 drosophila
118	16	61.5	181	2	06ZTR4_HUMAN	06ztr4 homo sapien	191	16	61.5	290	2	04EX06_LEITMA	04ex06 leishmania
119	16	61.5	188	2	04RZL3_TETNG	04rzl3 tetraodon n	192	16	61.5	290	2	06RHX7_LYCES	06rhx7 lycopersico
120	16	61.5	189	2	05GXC7_XANOR	05gxc7 xanthomonas	193	16	61.5	291	2	05B3N4_EBENT	05b3n4 aspergillus
121	16	61.5	191	2	0700R6_GIALA	0700r6 giardia lam	194	16	61.5	294	2	06Z0N2_CABER	06z0n2 caenorhabdi
122	16	61.5	192	2	080A8_TOSV	080a8 toscana vir	195	16	61.5	294	2	09XUS0_CABEL	09xus0 caenorhabdi
123	16	61.5	203	2	06ZVU4_HUMAN	06zv4 homo sapien	196	16	61.5	294	2	04RJG2_TETNG	04rjg2 tetraodon n
124	16	61.5	206	2	04SKY8_TETNG	04sky8 tetraodon n	197	16	61.5	296	2	07X9B9_ARATH	07x9b9 arabidopsis
125	16	61.5	213	2	074ZG8_ASHGO	074zg8 ashbya goss	198	16	61.5	298	2	08H5F5_ORYSA	08h5f5 oryza sativ
126	16	61.5	214	2	09W3F3_DROME	09w3f3 drosophila	199	16	61.5	300	2	04SUT2_TETNG	04sut2 tetraodon n
127	16	61.5	215	2	07R428_GIALA	07r428 giardia lam	200	16	61.5	306	2	084566_CHLTR	084566 chlamydia t
128	16	61.5	215	2	060UH4_9GAMM	060uh4 methylomona	201	16	61.5	309	2	04STZ8_TETNG	04stz8 tetraodon n
129	16	61.5	215	2	060UH6_9GAMM	060uh6 methylomona	202	16	61.5	311	2	051PB5_MAGGR	051pb5 magnaporthe
130	16	61.5	216	2	05TV10_ANOGA	05tv10 anopheles g	203	16	61.5	317	1	WRK15_ARATH	022176 arabidopsis
131	16	61.5	218	2	08G4A2_BIFLO	08g4a2 blitidobacte	204	16	61.5	321	2	09ZUS6_ARATH	09zun8 arabidopsis
132	16	61.5	219	2	063SB2_BURPS	063sb2 burkholderi	205	16	61.5	335	2	06AU3_ORYSA	06au3 oryza sativ
133	16	61.5	221	2	05WNK7_CABER	05wnk7 caenorhabdi	206	16	61.5	335	2	0727E2_DESVH	0727e2 desulfovibr
134	16	61.5	222	2	04VY8_DROME	04vy8 drosophila	207	16	61.5	340	2	07X1L8_ORYSA	07x1l8 oryza sativ
135	16	61.5	225	2	027013_TENMO	027013 tenebrio mo	208	16	61.5	346	2	07ZP51_9H1V1	07zp51 human immun
136	16	61.5	225	2	027013_TENMO	027013 tenebrio mo	209	16	61.5	346	2	04TW08_9H1V1	04tw08 human immun
137	16	61.5	228	2	04V5S4_DROME	04v5s4 drosophila	210	16	61.5	347	2	04QC08_LEITMA	04qc08 leishmania
138	16	61.5	232	2	062LR3_BURMA	062lr3 burkholderi	211	16	61.5	349	1	PAK4_PAT	088336 rattus nov
139	16	61.5	233	2	08LHK2_ORYSA	08lhk2 oryza sativ	212	16	61.5	349	2	04TW05_9H1V1	04tw05 human immun
140	16	61.5	237	2	06YU04_ORYSA	06yu04 oryza sativ	213	16	61.5	351	2	04TW09_9H1V1	04tw09 human immun
141	16	61.5	238	2	070U60_GIALA	070u60 giardia lam	214	16	61.5	353	2	067VP8_ORYSA	067vp8 oryza sativ
142	16	61.5	240	2	093877_FUSOX	093877 fusarium ox	215	16	61.5	353	2	04TWU7_9H1V1	04twu7 human immun
143	16	61.5	241	2	09VVW4_DROME	09vvw4 drosophila	216	16	61.5	354	2	04TWU7_9H1V1	04twu7 human immun
144	16	61.5	242	2	004701_NECCHA	004701 nectria hae	217	16	61.5	355	2	08M1T0_BOVIN	08m1t0 bos taurus
145	16	61.5	243	2	05QHS3_BRABE	05qhs3 brachydanio	218	16	61.5	358	2	04TWV2_9H1V1	04twv2 human immun
146	16	61.5	245	1	CTRB_GADMO	05046 gadus morhu	219	16	61.5	358	2	053XJ7_ORYSA	053xj7 oryza sativ
147	16	61.5	247	2	018458_HETGL	018458 heterodera	220	16	61.5	362	2	08CSN9_MOUSE	08csn9 mus musculu
148	16	61.5	249	2	05VRX0_ORYSA	05vrx0 oryza sativ	221	16	61.5	363	2	05H1J8_XANOR	05h1j8 xanthomonas
149	16	61.5	251	2	024774_ENCBU	024774 encyrtus	222	16	61.5	366	2	05LUJ3_STIPO	05luj3 silicibacte
150	16	61.5	251	2	070LQ4_ENCBU	070lq4 encyrtus	223	16	61.5	370	2	07XEN8_ORYSA	07xen8 oryza sativ
151	16	61.5	252	1	TRYI_DROME	P52905 drosophila	224	16	61.5	371	2	04P4V1_USTMA	04p4v1 usciago ma
152	16	61.5	252	2	0540Z6_DROME	0540z6 drosophila	225	16	61.5	373	2	06ABK7_PROAC	06abk7 propionibac
153	16	61.5	254	2	060K89_CABER	060k89 caenorhabdi	226	16	61.5	373	2	05SNS5_BRABE	05sns5 brachydanio
154	16	61.5	258	2	09Y0B9_DROME	09y0b9 drosophila	227	16	61.5	374	2	07PHB7_ANOGA	07phb7 caenorhabdi
155	16	61.5	259	2	09GQ40_GIALA	09gq40 giardia lam	228	16	61.5	374	2	0721R1_CABEL	0721r1 caenorhabdi
156	16	61.5	259	2	058PZ0_ARATH	058p20 arabidopsis	229	16	61.5	375	2	094699_PABPR	094699 paramectum
157	16	61.5	259	2	08H1N1_ARATH	08h1n1 arabidopsis	230	16	61.5	376	2	064CR1_9ARCH	064cr1 uncultured
158	16	61.5	259	2	06ACW2_BRABE	06acw2 brachydanio	231	16	61.5	377	2	057MD2_9REYP	057md2 trypanosoma
159	16	61.5	260	2	06ZPV9_XENTR	06zpv9 xenopus tro	232	16	61.5	378	2	063RGO_BURPS	063rgo burkholderi
160	16	61.5	261	2	09W7Q3_PAROL	09w7q3 paracititly	233	16	61.5	378	2	062166_BURMA	062166 burkholderi
161	16	61.5	261	2	06DHD9_BRABE	06dhd9 canine ente	234	16	61.5	380	2	062166_BURMA	062166 burkholderi
162	16	61.5	262	1	VMEI_CVCAI	07z113 human immun	235	16	61.5	380	2	0970G3_ARATH	0970g3 arabidopsis
163	16	61.5	262	1	053KW3_ORYSA	053kw3 oryza sativ	236	16	61.5	382	2	08XN74_CUOPE	08xn74 clostridium
164	16	61.5	263	1	CTRA_GADMO	P47796 gadus morhu	237	16	61.5	382	2	04QBZ6_LEITMA	04qbz6 leishmania
165	16	61.5	263	1	09PWQ6_GADMO	09pw6 gadus morhu	238	16	61.5	388	2	06P6D1_HUMAN	06p6d1 homo sapien
166	16	61.5	263	2	04SEW1_TETNG	04sew1 tetraodon n	239	16	61.5	392	2	09D646_MOUSE	09d646 mus musculu
167	16	61.5	263	2	05H2D0_XENTR	05h2d0 xenopus tro	240	16	61.5	394	2	09GQ4_IALA	09gq47 giardia lam
168	16	61.5	263	2	051029_XENTR	051029 xenopus tro	241	16	61.5	397	2	0623N1_CABER	0623n1 caenorhabdi
169	16	61.5	263	2	06GNF7_XENTLA	06gnf7 xenopus lae	242	16	61.5	397	2	09XUB6_CABEL	09xub6 caenorhabdi
170	16	61.5	263	2	06PSS4_XENTLA	06pss4 xenopus lae	243	16	61.5	399	2	07SET6_NEURCA	07set6 neurospora
171	16	61.5	263	2	07S1S4_XENTLA	07s1s4 xenopus lae	244	16	61.5	400	2	063R93_BURPS	063r93 burkholderi
172	16	61.5	264	2	086EC4_SCHUA	086ec4 schistosoma	245	16	61.5	401	2	07PV94_ANOGA	07pv94 anopheles g
173	16	61.5	264	2	04QY78_SPAUV	04qy78 sparus aura	246	16	61.5	404	2	09C211_NEURCA	09c211 neurospora
174	16	61.5	264	2	05GIU4_BRABE	05giu4 brachydanio	247	16	61.5	404	2	04RPND_TETNG	04rpnd tetraodon n
175	16	61.5	266	2	05MS69_LEGPA	05ms69 legionella	248	16	61.5	410	2	05H2U4_XENTR	05h2u4 xenopus tro
176	16	61.5	271	2	04S096_TETNG	04s096 tetraodon n	249	16	61.5	411	2	095SV1_DROME	095sv1 drosophila
177	16	61.5	271	2			250	16	61.5				



251	16	61.5	411	2	Q9VWV5_DROME	Q9VWV5_drosophila	324	16	61.5	572	2	Q42728_RHO70	Q42728_rhodospirid
252	16	61.5	415	2	Q5KC03_CRYNE	Q5KC03_cryptococcu	325	16	61.5	574	1	GAGJ_DROFY	P21331_drosophila
253	16	61.5	415	2	Q5TVP2_ANOGA	Q5TVP2_anopheles g	326	16	61.5	574	2	Q5X7Q4_BOTCI	Q5X7Q4_botrytis ci
254	16	61.5	415	2	Q7V960_PROMM	Q7V960_prochloroco	327	16	61.5	574	2	Q4RELI_MACPA	Q4RELI_maccia fasc
255	16	61.5	417	2	Q4Q802_LEIMA	Q4Q802_leishmania	328	16	61.5	581	2	Q96WM9_BOTCI	Q96WM9_botrytis ci
256	16	61.5	420	2	Q4QD22_LEIMA	Q4QD22_leishmania	329	16	61.5	587	2	Q7R120_GIALA	Q7R120_giardia lam
257	16	61.5	421	2	Q6W340_GHIV1	Q6W340_human immun	330	16	61.5	589	2	Q6GQU6_MOUSE	Q6GQU6_mus muscucu
258	16	61.5	422	1	RAD52_CHICK	P39022_gallus gall	331	16	61.5	590	1	IRAK2_HUMAN	Q43187_homo sapien
259	16	61.5	426	2	Q72W52_LEPIC	Q72W52_leptospiro	332	16	61.5	600	1	EGFL5_HUMAN	Q9H144_homo sapien
260	16	61.5	426	2	Q8P9T9_LEPIN	Q8P9T9_leptospiro	333	16	61.5	603	2	Q585K6_9TRYR	Q585K6_hypanosoma
261	16	61.5	430	2	Q840C2_9GAMM	Q840C2_cellulibrio	334	16	61.5	603	2	Q4UHG3_9LAMI	Q4UHG3_pellilla cit
262	16	61.5	436	2	Q26859_TRYCO	Q26859_trypanosoma	335	16	61.5	607	2	Q5U388_BRARE	Q5U388_brychanidantio
263	16	61.5	440	2	Q941N4_AVEBA	Q941N4_avena sativ	336	16	61.5	608	2	Q57UW7_9TRYR	Q57UW7_9trypanosoma
264	16	61.5	441	2	Q6UAP5_TETNG	Q6UAP5_tetradodon n	337	16	61.5	613	2	Q957Q2_DROME	Q957Q2_drosophila
265	16	61.5	443	2	Q6AZB6_BRARE	Q6AZB6_brychanidantio	338	16	61.5	613	2	Q9V5F6_DROME	Q9V5F6_drosophila
266	16	61.5	446	2	Q7XGP0_ORISA	Q7XGP0_oryza sativ	339	16	61.5	616	2	Q738S5_BACCI	Q738S5_bacillus ce
267	16	61.5	446	2	Q8SB64_ORISA	Q8SB64_oryza sativ	340	16	61.5	622	2	Q7QMD2_GIALA	Q7QMD2_giardia lam
268	16	61.5	447	2	Q8TL45_METAC	Q8TL45_methanosarc	341	16	61.5	625	2	Q5K546_HUMAN	Q5K546_homo sapien
269	16	61.5	448	2	Q5GVX5_XANOR	Q5GVX5_xanthomonas	342	16	61.5	625	2	Q5R810_PONPY	Q5R810_pongo pygma
270	16	61.5	450	2	Q7XP30_ORISA	Q7XP30_oryza sativ	343	16	61.5	629	2	Q9DWH3_RCMVM	Q9DWH3_rat cytoomeg
271	16	61.5	452	2	Q8O000_MEIMA	Q8O000_methanosarc	344	16	61.5	632	2	Q6YXW2_ORISA	Q6YXW2_oryza sativ
272	16	61.5	453	2	Q20490_CABEL	Q20490_caenorhabdi	345	16	61.5	633	2	Q8WTK1_9HIVA	Q8WTK1_meloidogyne
273	16	61.5	455	2	Q4PT10_ARATH	Q4PT10_arabidopsis	346	16	61.5	636	2	Q61BP2_CABBR	Q61BP2_caenorhabdi
274	16	61.5	455	2	Q6IFV6_RAT	Q6IFV6_rattus norv	347	16	61.5	638	1	YT32_CABEL	Q20680_caenorhabdi
275	16	61.5	457	2	Q9YQ21_9HERP	Q9YQ21_randid herpe	348	16	61.5	639	2	Q84K89_AVEBA	Q84K89_avena sativ
276	16	61.5	461	2	Q4Q6C9_LEIMA	Q4Q6C9_leishmania	349	16	61.5	642	2	Q9PPI1_CAMJE	Q9PPI1_campylobact
277	16	61.5	463	2	Q6Z3D6_ORISA	Q6Z3D6_oryza sativ	350	16	61.5	645	2	Q97448_GIALA	Q97448_giardia lam
278	16	61.5	464	2	Q5CTB1_CRYPV	Q5CTB1_cryptospori	351	16	61.5	650	2	Q4S9Q3_TETNG	Q4S9Q3_tetradodon n
279	16	61.5	464	2	Q5CIV7_CRYHO	Q5CIV7_cryptospori	352	16	61.5	652	2	Q7QY47_GIALA	Q7QY47_giardia lam
280	16	61.5	466	2	Q7QVC3_GIALA	Q7QVC3_giardia lam	353	16	61.5	660	2	Q6PCHO_XENIA	Q6PCHO_xenopus lae
281	16	61.5	466	2	Q5UBX1_9HIV1	Q5UBX1_human immun	354	16	61.5	665	1	YDER_SCHPO	Q10447_schizosacch
282	16	61.5	468	2	Q8GY60_ARATH	Q8GY60_arabidopsis	355	16	61.5	667	2	Q9XTK3_GIALA	Q9XTK3_giardia lam
283	16	61.5	470	2	Q8ZBG0_STRAM	Q8ZBG0_streptomyce	356	16	61.5	679	2	Q6PUC7_9HIVA	Q6PUC7_homo sapien
284	16	61.5	471	2	Q5TYV3_MOCRA	Q5TYV3_nocardia fa	357	16	61.5	681	2	Q6LBN7_BOVIN	Q6LBN7_bos taurus
285	16	61.5	473	2	Q5ZCJ2_MAGGR	Q5ZCJ2_magnaporthe	358	16	61.5	687	2	Q817G1_CABEL	Q817G1_caenorhabdi
286	16	61.5	473	2	Q5FAD8_NEIG1	Q5FAD8_neisseria g	359	16	61.5	695	1	TRFL_HORSE	Q7R811_equus cabal
287	16	61.5	473	2	Q9JWV9_NEIMA	Q9JWV9_neisseria m	360	16	61.5	696	2	Q5R2V7_HUMAN	Q5R2V7_homo sapien
288	16	61.5	473	2	Q9JXZ7_NEIMB	Q9JXZ7_neisseria m	361	16	61.5	699	2	Q93HN9_STRAM	Q93HN9_streptomyce
289	16	61.5	475	2	Q9X8J5_STRCO	Q9X8J5_streptomyce	362	16	61.5	708	1	TRFL_BOVIN	Q24627_bos taurus
290	16	61.5	479	2	Q4SIL6_TETNG	Q4SIL6_tetradodon n	363	16	61.5	708	1	TRFL_BUBRU	Q77698_bubalus bub
291	16	61.5	480	2	Q7SZR0_BRARE	Q7SZR0_brychanidantio	364	16	61.5	708	1	TRFL_CAMDR	Q97un00_camelus dro
292	16	61.5	481	2	Q9D635_MOUSE	Q9D635_mus muscucu	365	16	61.5	708	1	TRFL_CAPHI	Q29477_capra hircu
293	16	61.5	486	2	Q9LQ21_ARATH	Q9LQ21_arabidopsis	366	16	61.5	708	2	Q6LEC7_BOVIN	Q6LEC7_bos taurus
294	16	61.5	488	2	Q5WMJ6_LEGPL	Q5WMJ6_legionella	367	16	61.5	711	1	TRFL_HUMAN	P02788_homo sapien
295	16	61.5	489	2	Q5AL61_CANAL	Q5AL61_candida alb	368	16	61.5	711	2	Q5EKS1_HUMAN	Q5EKS1_homo sapien
296	16	61.5	496	2	Q4NXY7_9DELT	Q4NXY7_aeromonas	369	16	61.5	713	2	Q5DSMO_HUMAN	Q5DSMO_homo sapien
297	16	61.5	501	2	Q4HVA7_GIBZE	Q4HVA7_gibberella	370	16	61.5	713	2	Q6FBS7_ACTAD	Q6FBS7_actinobact
298	16	61.5	510	1	MS2P_CRIGR	Q54862_cricetulus	371	16	61.5	719	1	Q757T5_ASHGO	Q757T5_ashya goss
299	16	61.5	513	2	Q4IVN9_AZOV1	Q4IVN9_azotobacter	372	16	61.5	722	2	Q757T5_ASHGO	Q757T5_ashya goss
300	16	61.5	517	2	Q4VBN3_RAT	Q4VBN3_rattus norv	373	16	61.5	722	2	Q6FBS2_CANCA	Q6FBS2_candida gla
301	16	61.5	520	2	Q8NCV0_HUMAN	Q8NCV0_homo sapien	374	16	61.5	734	2	Q9NSE6_HUMAN	Q9NSE6_homo sapien
302	16	61.5	520	2	Q8IYS4_HUMAN	Q8IYS4_homo sapien	375	16	61.5	740	2	Q9UPP2_HUMAN	Q9UPP2_homo sapien
303	16	61.5	522	2	Q53P11_ORISA	Q53P11_oryza sativ	376	16	61.5	741	2	Q57VP8_9TRYR	Q57VP8_9trypanosoma
304	16	61.5	523	2	Q86MK1_DROME	Q86MK1_drosophila	377	16	61.5	743	2	Q4QD80_LEIMA	Q4QD80_leishmania
305	16	61.5	528	2	Q53NH3_ORISA	Q53NH3_oryza sativ	378	16	61.5	777	2	Q4QD78_LEIMA	Q4QD78_leishmania
306	16	61.5	532	2	Q8XIC0_USITMA	Q8XIC0_ustilago ma	379	16	61.5	777	2	Q4QD78_LEIMA	Q4QD78_leishmania
307	16	61.5	533	2	Q4PEV3_USITMA	Q4PEV3_ustilago ma	380	16	61.5	783	2	Q9FEV4_STRCO	Q9FEV4_streptomyce
308	16	61.5	533	2	Q74EC9_GEOSL	Q74EC9_geobacter s	381	16	61.5	788	2	Q76128_9VIRU	Q76128_helicobasid
309	16	61.5	536	2	Q4U3U6_NEUAF	Q4U3U6_neutrospora	382	16	61.5	795	2	Q5NM24_9ARCH	Q5NM24_uncultured
310	16	61.5	536	2	Q4U3U4_9PEZ1	Q4U3U4_neutrospora	383	16	61.5	802	2	Q4IEN4_GIBZE	Q4IEN4_gibberella
311	16	61.5	537	1	Q4T8X5_TETNG	Q4T8X5_tetradodon n	384	16	61.5	810	2	Q4ID36_GIBZE	Q4ID36_gibberella
312	16	61.5	537	1	QAY_NEUCR	P1666_neutrospora	385	16	61.5	824	2	Q5VQ42_ORISA	Q5VQ42_oryza sativ
313	16	61.5	538	2	Q5BAQ4_EMENT	P1666_aspergillus	386	16	61.5	825	2	Q4Q608_LEIMA	Q4Q608_leishmania
314	16	61.5	542	2	Q6ZIB7_ORISA	Q6ZIB7_oryza sativ	387	16	61.5	828	2	Q7RPF7_PLAYO	Q7RPF7_plasmodium
315	16	61.5	542	2	Q9FP95_ARATH	Q9FP95_arabidopsis	388	16	61.5	842	2	Q61IW9_CABBR	Q61IW9_caenorhabdi
316	16	61.5	545	2	Q86ZM8_PODAP	Q86ZM8_podospora a	389	16	61.5	845	2	Q8P124_XANAC	Q8P124_xanthomonas
317	16	61.5	545	2	Q4IAQ1_GIBZE	Q4IAQ1_gibberella	390	16	61.5	845	2	Q8AK00_9HIV1	Q8AK00_human immun
318	16	61.5	546	2	Q5M281_9AGAR	Q5M281_pleurocous s	391	16	61.5	845	2	Q4Q607_LEIMA	Q4Q607_leishmania
319	16	61.5	549	2	Q51YV6_MAGGR	Q51YV6_magnaporthe	392	16	61.5	847	2	Q4Q609_LEIMA	Q4Q609_leishmania
320	16	61.5	552	2	Q4IRO1_GIBZE	Q4IRO1_gibberella	393	16	61.5	847	2	Q8AP19_9HIV1	Q8AP19_human immun
321	16	61.5	560	2	Q6DUJ9_XENTR	Q6DUJ9_xenopus tro	394	16	61.5	848	2	Q4QRT9_9HIV1	Q4QRT9_human immun
322	16	61.5	561	2	Q5BS14_EMENT	Q5BS14_aspergillus	395	16	61.5	851	2	Q4QWS5_9HIV1	Q4QWS5_human immun
323	16	61.5	561	2	Q96WN0_BOTCI	Q96WN0_botrytis ci	396	16	61.5	852	2	Q7RVV6_NEUCR	Q7RVV6_neutrospora

397	16	61.5	852	2	Q9HB44_NEUCR	Q9h44	neuropora	470	16	61.5	2128	2	Q5ONG8_ENTHI	Q5ong8	entamoeba h
398	16	61.5	852	2	Q70010_9H1V1	Q70010	human immun	471	16	61.5	2233	2	Q94711_PARTE	Q94711	paramecium
399	16	61.5	854	2	Q5BEC9_9H1V1	Q5bec9	human immun	472	16	61.5	2273	2	Q5ICB9_MACCP	Q5icb9	macaca faec
400	16	61.5	858	2	Q8AB09_9H1V1	Q8ae09	human immun	473	16	61.5	2377	2	Q7MQ80_BORBR	Q7mq80	bordelella
401	16	61.5	859	2	Q4Q027_LEI1MA	Q4q027	leishmania	474	16	61.5	2395	2	Q27167_PARTE	Q27167	paramecium
402	16	61.5	859	2	Q8ADK7_9H1V1	Q8ack7	human immun	475	16	61.5	2397	2	Q6Q7Y4_PARPR	Q6q7y4	paramecium
403	16	61.5	859	2	Q8UMPE_9H1V1	Q8ump8	human immun	476	16	61.5	2533	2	P90589_PARTE	P90589	paramecium
404	16	61.5	860	2	Q998E5_9H1V1	Q998e5	human immun	477	16	61.5	2533	2	Q27183_PARTE	Q27183	paramecium
405	16	61.5	864	2	Q4U536_9H1V1	Q4U536	human immun	478	16	61.5	2543	2	P90649_PARPR	P90649	paramecium
406	16	61.5	865	2	Q73342_9H1V1	Q73342	human immun	479	16	61.5	2715	1	G156_PARPR	G156	PARPR
407	16	61.5	865	2	Q73353_9H1V1	Q73353	human immun	480	16	61.5	2986	1	Q4HMT4_GIBBZ	Q4hmt4	gibberella
408	16	61.5	870	2	Q5EBD7_9H1V1	Q5ebd7	human immun	481	16	61.5	3596	1	RPOA_SHPV	RPOA	SHPV
409	16	61.5	877	2	Q5VG71_9H1V1	Q5vg71	human immun	482	16	61.5	3869	2	Q5CRG0_CRYPV	Q5crg0	cryptospori
410	16	61.5	879	2	Q4M9A0_ASPFU	Q4m9a0	aspergillus	483	16	61.5	3869	2	Q8EPQ3_CRYPV	Q8epq3	cryptospori
411	16	61.5	879	2	Q4PEC4_USTMA	Q4pec4	ustilago ma	484	16	61.5	4382	2	Q5DIC6_CVHSA	Q5dic6	cvhsa
412	16	61.5	881	2	Q4N476_TBHPA	Q4n476	thelateria p	485	16	61.5	4382	2	Q5Y188_CVHSA	Q5y188	cvhsa
413	16	61.5	884	2	Q8Q718_9H1V1	Q8q718	human immun	486	16	61.5	4382	2	Q692B5_CVHSA	Q692b5	cvhsa
414	16	61.5	889	2	Q8IUQ3_PLAF7	Q8i1q3	plasmodium	487	16	61.5	4382	2	Q6JH33_CVHSA	Q6jh33	cvhsa
415	16	61.5	892	2	Q5SSQ4_CRYNE	Q5sfq4	cryptococcu	488	16	61.5	4382	2	Q6JH47_CVHSA	Q6jh47	cvhsa
416	16	61.5	897	2	Q5KGF6_CRYNE	Q5kff6	cryptococcu	489	16	61.5	4382	2	Q6R7Y7_CVHSA	Q6r7y7	cvhsa
417	16	61.5	907	2	Q5FM49_ARYTH	Q5fm49	arabidopsis	490	16	61.5	4382	2	Q6RCM6_CVHSA	Q6rcm6	cvhsa
418	16	61.5	911	2	Q80TU8_MOUSE	Q80tj8	mus musculu	491	16	61.5	4382	2	Q6RCX7_CVHSA	Q6rcx7	cvhsa
419	16	61.5	912	2	Q4JHL4_FUGRU	Q4jhl4	fugu rubrip	492	16	61.5	4382	2	Q6RCY8_CVHSA	Q6rcy8	cvhsa
420	16	61.5	918	2	Q6E7C9_9URUC	Q6e7c9	ikopileura	493	16	61.5	4382	2	Q6RCZ9_CVHSA	Q6rcz9	cvhsa
421	16	61.5	920	2	Q7XGN4_ORYSA	Q7xgn4	oryza sativ	494	16	61.5	4382	2	Q6RD10_CVHSA	Q6rd10	cvhsa
422	16	61.5	920	2	Q94GU3_ORYSA	Q94gu3	oryza sativ	495	16	61.5	4382	2	Q6RD21_CVHSA	Q6rd21	cvhsa
423	16	61.5	929	2	Q5SED8_9H1V1	Q5sed8	human immun	496	16	61.5	4382	2	Q6RD33_CVHSA	Q6rd33	cvhsa
424	16	61.5	934	2	Q87R90_VIBPA	Q87r90	vibrio para	497	16	61.5	4382	2	Q6RD43_CVHSA	Q6rd43	cvhsa
425	16	61.5	935	2	Q4P9H9_USTMA	Q4p9h9	ustilago ma	498	16	61.5	4382	2	Q6RD54_CVHSA	Q6rd54	cvhsa
426	16	61.5	939	2	Q5RHW8_BRARE	Q5rhw8	brachyana	499	16	61.5	4382	2	Q6RD65_CVHSA	Q6rd65	cvhsa
427	16	61.5	967	2	Q4T490_TETNG	Q4t490	tetradion n	500	16	61.5	4382	2	Q6VA79_CVHSA	Q6va79	cvhsa
428	16	61.5	975	2	P91357_CABEL	P91357	caenorhabd	501	16	61.5	4382	2	Q6VA90_CVHSA	Q6va90	cvhsa
429	16	61.5	999	1	SMG_DROME	Q23372	drosophila	502	16	61.5	4382	2	Q6VYA1_CVHSA	Q6vya1	cvhsa
430	16	61.5	1031	2	Q65500_ARYTH	Q65500	arabidopsis	503	16	61.5	4382	2	Q6WGN0_CVHSA	Q6wgn0	cvhsa
431	16	61.5	1038	2	Q583N6_9TRYP	Q583n6	trypanosoma	504	16	61.5	5560	1	SPEN_DROME	SPEN	DROME
432	16	61.5	1039	2	Q4P295_USTMA	Q4p295	ustilago ma	505	16	61.5	6880	2	Q6S8D8_CVHSA	Q6s8d8	cvhsa
433	16	61.5	1069	2	Q592X1_CANAL	Q592x1	candida alb	506	16	61.5	7073	1	RIAB_CVHSA	RIAB	CVHSA
434	16	61.5	1069	2	Q9U3F2_CABEL	Q9uf2	caenorhabd	507	16	61.5	7073	2	Q5DIC7_CVHSA	Q5dic7	cvhsa
435	16	61.5	1071	2	Q61ZG1_CABER	Q61zgi	caenorhabd	508	16	61.5	7073	2	Q5Y189_CVHSA	Q5y189	cvhsa
436	16	61.5	1105	2	Q20371_CABEL	Q20371	caenorhabd	509	16	61.5	7073	2	Q6Y2B6_CVHSA	Q6y2b6	cvhsa
437	16	61.5	1136	2	Q92C14_LISIN	Q92c14	liisteria in	510	16	61.5	7073	2	Q6JH40_CVHSA	Q6jh40	cvhsa
438	16	61.5	1154	2	Q9GQ46_GIALA	Q9gq46	giardia lam	511	16	61.5	7073	2	Q6JH48_CVHSA	Q6jh48	cvhsa
439	16	61.5	1168	2	Q6AVU1_ORYSA	Q6av11	oryza sativ	512	16	61.5	7073	2	Q6R7Y8_CVHSA	Q6r7y8	cvhsa
440	16	61.5	1191	2	Q8TOL3_DROME	Q8tol3	drosophila	513	16	61.5	7073	2	Q6RPE8_CVHSA	Q6rpe8	cvhsa
441	16	61.5	1194	2	Q4EX33_LEIMA	Q4fx33	leishmania	514	16	61.5	7073	2	Q6UZP1_CVHSA	Q6uzp1	cvhsa
442	16	61.5	1194	2	Q76M68_RAT	Q76m68	rattus norv	515	16	61.5	7073	2	Q6VZB5_CVHSA	Q6vzb5	cvhsa
443	16	61.5	1227	2	Q6IFU2_SCHMA	Q6ifu2	schistosoma	516	16	61.5	7073	2	Q6V5B6_CVHSA	Q6v5b6	cvhsa
444	16	61.5	1243	2	Q8QTA3_WSSV	Q8qta3	white spot	517	16	61.5	7073	2	Q6RCW7_CVHSA	Q6rcw7	cvhsa
445	16	61.5	1243	2	Q8VAP5_WSSV	Q8vap5	white spot	518	16	61.5	7073	2	Q6RCX8_CVHSA	Q6rcx8	cvhsa
446	16	61.5	1243	2	Q91LM1_WSSV	Q91lm1	white spot	519	16	61.5	7073	2	Q6RCY9_CVHSA	Q6rcy9	cvhsa
447	16	61.5	1270	2	Q5UXR6_HUMAN	Q5uxr6	homo sapien	520	16	61.5	7073	2	Q6RD00_CVHSA	Q6rd00	cvhsa
448	16	61.5	1310	1	GPI25_MOUSE	Q7tc36	mus musculu	521	16	61.5	7073	2	Q6RD11_CVHSA	Q6rd11	cvhsa
449	16	61.5	1318	2	Q4QDAS_LEIMA	Q4qdas	leishmania	522	16	61.5	7073	2	Q6RD22_CVHSA	Q6rd22	cvhsa
450	16	61.5	1332	2	Q4FPD5_LEIMA	Q4fpd5	leishmania	523	16	61.5	7073	2	Q6RD33_CVHSA	Q6rd33	cvhsa
451	16	61.5	1353	2	Q4QBR1_LEIMA	Q4qbr1	leishmania	524	16	61.5	7073	2	Q6RD44_CVHSA	Q6rd44	cvhsa
452	16	61.5	1363	2	Q4LIW7_9BURK	Q4liw7	burkholderi	525	16	61.5	7073	2	Q6RD55_CVHSA	Q6rd55	cvhsa
453	16	61.5	1367	2	Q4QAT6_LEIMA	Q4qat6	leishmania	526	16	61.5	7073	2	Q6RD66_CVHSA	Q6rd66	cvhsa
454	16	61.5	1369	2	Q4QAT7_LEIMA	Q4qat7	leishmania	527	16	61.5	7073	2	Q6VA80_CVHSA	Q6va80	cvhsa
455	16	61.5	1380	1	CYAA_LEIDO	Q27675	leishmania	528	16	61.5	7073	2	Q6VAY1_CVHSA	Q6vay1	cvhsa
456	16	61.5	1467	2	Q61QNF_CABER	Q61qnf	caenorhabd	529	16	61.5	7073	2	Q6VYA2_CVHSA	Q6vya2	cvhsa
457	16	61.5	1474	2	Q625Q4_CABER	Q625q4	caenorhabd	530	16	61.5	7073	2	Q4UDQ5_CVHSA	Q4udq5	cvhsa
458	16	61.5	1486	2	Q4SDM0_TETNG	Q4sdm0	tetradion n	531	16	61.5	7073	2	Q4UDR8_CVHSA	Q4udr8	cvhsa
459	16	61.5	1557	2	Q4QHE6_LEIMA	Q4qhe6	leishmania	532	16	61.5	7073	2	Q93940_PODAN	Q93940	podospora a
460	16	61.5	1558	2	Q55TP3_CRYNE	Q55tp3	cryptococcu	533	15	57.7	33	2	Q9BY80_HUMAN	Q9by80	homo sapien
461	16	61.5	1585	2	Q5KIP2_CRYNE	Q5kip2	cryptococcu	534	15	57.7	40	1	MT1_DROME	MT1	DROME
462	16	61.5	1769	2	Q4QFD0_LEIMA	Q4qfd0	leishmania	535	15	57.7	40	1	MTI_DROME	MTI	DROME
463	16	61.5	1818	1	Q7WC76_BORPA	Q7wc76	bordelella	536	15	57.7	40	1	COAT2_BROST	COAT2	BROST
464	16	61.5	1823	1	ENC_DROME	Q8mx1	drosophila	537	15	57.7	43	1	COAT2_XANCP	COAT2	XANCP
465	16	61.5	1823	1	Q5BTH9_DROME	Q5bth9	drosophila	538	15	57.7	43	1	Q4UVU8_XANCP	Q4uvu8	xanthomonas
466	16	61.5	1823	1	Q7KV61_DROME	Q7kv61	drosophila	539	15	57.7	43	2	Q4UVU8_XANCP	Q4uvu8	xanthomonas
467	16	61.5	2071	2	Q7QV58_GIALA	Q7qv58	giardia lam	540	15	57.7	46	2	Q4UZU3_XANCP	Q4uzu3	xanthomonas
468	16	61.5	2087	2	Q4S488_TETNG	Q4s488	tetradion n	541	15	57.7	46	2	Q8PDK1_XANCP	Q8pdk1	xanthomonas
469	16	61.5	2117	2	Q4QIS6_LEIMA	Q4qis6	leishmania	542	15	57.7	46	2			

543	15	57.7	53	2	Q69338_9GAMA	Q69338	salixirine	616	15	57.7	101	2	009129_MOUSE	009129	mus musculus
544	15	57.7	54	2	Q74959_GEOSL	Q74959	geobacter s	617	15	57.7	101	2	Q71260_9H1V1	Q71260	human immun
545	15	57.7	56	2	Q5C2S7_SCHJA	Q5C2S7	schistosoma	618	15	57.7	101	2	Q90DH4_9H1V1	Q90DH4	human immun
546	15	57.7	57	2	Q6OCB8_MERCA	Q6OCB8	methylcocc	619	15	57.7	101	2	Q90DH7_9H1V1	Q90DH7	human immun
547	15	57.7	58	2	Q96814_9BETA	Q96814	human herpe	620	15	57.7	102	1	VP07_BPAB8	VP07	bacterioph
548	15	57.7	58	2	Q96915_9BETA	Q96915	human herpe	621	15	57.7	102	2	Q6MN21_BDEBA	Q6MN21	bdellovibri
549	15	57.7	58	2	Q4TUC8_TETNG	Q4TUC8	tetradon n	622	15	57.7	102	2	Q71271_9H1V1	Q71271	human immun
550	15	57.7	59	2	Q7S2A3_NEUCR	Q7S2A3	neutropora	623	15	57.7	102	2	Q90D14_9H1V1	Q90D14	human immun
551	15	57.7	59	2	Q5C1N4_SCHJA	Q5C1N4	schistosoma	624	15	57.7	103	2	Q70RX6_GIALA	Q70RX6	giardia lam
552	15	57.7	59	2	Q4FP36_TETNG	Q4FP36	tetradon n	625	15	57.7	103	2	Q9U0T5_LOGMT	Q9U0T5	locustia mig
553	15	57.7	61	2	Q86TX3_HUMAN	Q86TX3	homo sapien	626	15	57.7	103	2	Q806T2_9HEPC	Q806T2	hepatitis c
554	15	57.7	63	2	Q6F0N6_CANGA	Q6F0N6	candida gla	627	15	57.7	104	2	Q7PG48_ANGCA	Q7PG48	anopheles g
555	15	57.7	64	2	Q6KOM3_9VIRU	Q6KOM3	human astro	628	15	57.7	104	2	Q69W53_ORYSA	Q69W53	oryza sativ
556	15	57.7	64	2	Q8JMX7_9VIRU	Q8JMX7	human astro	629	15	57.7	104	2	Q84BA7_ERVCH	Q84BA7	eryincha chr
557	15	57.7	67	2	Q4TGES_TETNG	Q4TGES	tetradon n	630	15	57.7	106	2	Q8N0X0_CORGL	Q8N0X0	corynebacte
558	15	57.7	68	2	Q4THB0_TETNG	Q4THB0	tetradon n	631	15	57.7	107	2	Q7R3Y4_GIALA	Q7R3Y4	giardia lam
559	15	57.7	72	2	Q5DEB6_ANGCA	Q5DEB6	anopheles g	632	15	57.7	107	2	Q84N43_9PRBA	Q84N43	zapoteca te
560	15	57.7	73	2	Q7R2J5_GIALA	Q7R2J5	giardia lam	633	15	57.7	107	2	Q9D2X3_9H1V1	Q9D2X3	human immun
561	15	57.7	73	2	Q5RL97_PIG	Q5RL97	sus scrofa	634	15	57.7	107	2	Q9D2X4_9H1V1	Q9D2X4	human immun
562	15	57.7	74	2	Q62L27_BURMA	Q62L27	burholderi	635	15	57.7	107	2	Q9D2X5_9H1V1	Q9D2X5	human immun
563	15	57.7	74	2	Q8YS55_ANNAP	Q8YS55	annap	636	15	57.7	107	2	Q9D2X6_9H1V1	Q9D2X6	human immun
564	15	57.7	75	2	Q5YR21_9VIRU	Q5YR21	rock bream	637	15	57.7	107	2	Q9D2V5_9H1V1	Q9D2V5	human immun
565	15	57.7	77	2	Q4UZT9_XANCP	Q4UZT9	xanthomonas	638	15	57.7	109	1	YG1C_YEAST	YG1C	yeast
566	15	57.7	77	2	Q8PDJ8_XANCP	Q8PDJ8	xanthomonas	639	15	57.7	109	2	Q8CEA7_MOUSE	Q8CEA7	mus musculus
567	15	57.7	78	2	Q06467_XANCA	Q06467	xanthomonas	640	15	57.7	109	2	Q8JG58_AMEME	Q8JG58	ambystoma m
568	15	57.7	79	2	Q26086_9TIRB	Q26086	polycelis n	641	15	57.7	111	2	Q6CTJ3_KULIA	Q6CTJ3	kluveromyc
569	15	57.7	79	2	Q92NT4_RHIME	Q92NT4	rhicobium m	642	15	57.7	111	2	Q9Z8B2_CHLPN	Q9Z8B2	chlamydia p
570	15	57.7	80	2	Q4J5G5_AZOV1	Q4J5G5	azotobacter	643	15	57.7	111	2	Q8C5D5_MOUSE	Q8C5D5	mus musculus
571	15	57.7	80	2	Q5NCE5_MOUSE	Q5NCE5	mus musculus	644	15	57.7	112	2	Q56714_9HEPC	Q56714	hepatitis c
572	15	57.7	80	2	Q69752_9H1V1	Q69752	human immun	645	15	57.7	112	2	Q56716_9HEPC	Q56716	hepatitis c
573	15	57.7	81	2	Q80226_9H1V1	Q80226	human immun	646	15	57.7	112	2	Q56717_9HEPC	Q56717	hepatitis c
574	15	57.7	82	1	LCR45_ARATH	LCR45	arabidopsis	647	15	57.7	112	2	Q56718_9HEPC	Q56718	hepatitis c
575	15	57.7	85	2	Q09BYT5_HUMAN	Q09BYT5	homo sapien	648	15	57.7	112	2	Q56719_9HEPC	Q56719	hepatitis c
576	15	57.7	85	2	Q84GQ2_9SYNE	Q84GQ2	synecococc	649	15	57.7	112	2	Q56724_9HEPC	Q56724	hepatitis c
577	15	57.7	85	2	P88625_9H1V1	P88625	human immun	650	15	57.7	112	2	Q56726_9HEPC	Q56726	hepatitis c
578	15	57.7	86	2	Q6TRY3_CULQU	Q6TRY3	culex quinq	651	15	57.7	112	2	Q56727_9HEPC	Q56727	hepatitis c
579	15	57.7	88	2	Q94203_CLAFO	Q94203	cladosporiu	652	15	57.7	112	2	Q56728_9HEPC	Q56728	hepatitis c
580	15	57.7	88	2	Q90DH6_9H1V1	Q90DH6	human immun	653	15	57.7	112	2	Q56729_9HEPC	Q56729	hepatitis c
581	15	57.7	89	2	Q43185_HUMAN	Q43185	homo sapien	654	15	57.7	112	2	Q56730_9HEPC	Q56730	hepatitis c
582	15	57.7	89	2	Q51BM5_ENZHI	Q51BM5	entomoba h	655	15	57.7	112	2	Q56731_9HEPC	Q56731	hepatitis c
583	15	57.7	89	2	Q41UR5_AZOV1	Q41UR5	azotobacter	656	15	57.7	112	2	Q90DH8_9H1V1	Q90DH8	human immun
584	15	57.7	90	2	Q6TVM4_9POXY	Q6TVM4	orf virus	657	15	57.7	113	1	Y011_BP74	Y011	bp74
585	15	57.7	91	2	Q4XC49_PLACH	Q4XC49	plasmidom	658	15	57.7	113	2	Q5A4Z9_CANAL	Q5A4Z9	candida alb
586	15	57.7	91	2	Q96Y16_ARATH	Q96Y16	arabidopsis	659	15	57.7	114	2	Q614J6_CABBR	Q614J6	caenorhabdi
587	15	57.7	91	2	Q39862_9VIRU	Q39862	human astro	660	15	57.7	114	2	Q95ZX8_CABEL	Q95ZX8	caenorhabdi
588	15	57.7	91	2	Q6TW04_9POXY	Q6TW04	orf virus	661	15	57.7	114	2	Q90D10_9H1V1	Q90D10	human immun
589	15	57.7	92	2	Q9UIG9_HUMAN	Q9UIG9	homo sapien	662	15	57.7	115	2	Q9S9A1_VISAL	Q9S9A1	viscum albu
590	15	57.7	92	2	Q8SAX8_ORYSA	Q8SAX8	oryza sativ	663	15	57.7	116	2	Q6RXO0_9H1V1	Q6RXO0	human immun
591	15	57.7	92	2	P88627_9H1V1	P88627	human immun	664	15	57.7	117	2	Q4NZF5_9DBLT	Q4NZF5	aeromonas
592	15	57.7	93	2	Q5FMFO_LACAC	Q5FMFO	lactobacill	665	15	57.7	117	2	Q8U006_9H1V1	Q8U006	human immun
593	15	57.7	93	2	Q41098_CHYPI	Q41098	paramecium	666	15	57.7	117	2	Q8U007_9H1V1	Q8U007	human immun
594	15	57.7	93	2	Q5FZJ8_9HEPC	Q5FZJ8	hepatitis c	667	15	57.7	117	2	Q8U008_9H1V1	Q8U008	human immun
595	15	57.7	93	2	Q5FZK1_9HEPC	Q5FZK1	hepatitis c	668	15	57.7	118	1	CG023_HUMAN	CG023	homo sapien
596	15	57.7	93	2	Q5FZK8_9HEPC	Q5FZK8	hepatitis c	669	15	57.7	118	2	Q8RFE3_FUSNN	Q8RFE3	fungus
597	15	57.7	93	2	Q78021_9H1V1	Q78021	human immun	670	15	57.7	119	2	Q7U993_SYMPX	Q7U993	synecococc
598	15	57.7	94	2	Q90DH5_9H1V1	Q90DH5	human immun	671	15	57.7	120	2	Q8U009_9H1V1	Q8U009	human immun
599	15	57.7	94	2	Q9PXN5_9H1V1	Q9PXN5	human immun	672	15	57.7	121	2	Q69960_9H1V1	Q69960	human immun
600	15	57.7	95	2	Q8C8S2_MOUSE	Q8C8S2	mus musculus	673	15	57.7	121	2	Q69964_9H1V1	Q69964	human immun
601	15	57.7	96	2	Q9BYU5_HUMAN	Q9BYU5	homo sapien	674	15	57.7	122	1	YK4_CVHSA	YK4	cvhsa
602	15	57.7	96	2	Q7YTR6_CABEL	Q7YTR6	caenorhabdi	675	15	57.7	122	2	Q5K0L2_ORYSA	Q5K0L2	oryza sativ
603	15	57.7	96	2	Q8VD96_MESAU	Q8VD96	mesocricetu	676	15	57.7	122	2	Q5X205_LEGPA	Q5X205	legionella
604	15	57.7	96	2	Q77367_9H1V1	Q77367	human immun	677	15	57.7	122	2	Q5K2J7_LBOPH	Q5K2J7	leishmania
605	15	57.7	96	2	Q90DH9_9H1V1	Q90DH9	human immun	678	15	57.7	122	2	Q82209_CHLCP	Q82209	chlamydia
606	15	57.7	97	2	Q8TFPH4_METAC	Q8TFPH4	methanosarc	679	15	57.7	122	2	Q5Y181_CVHBA	Q5Y181	sars corona
607	15	57.7	97	2	Q4Y107_PLABE	Q4Y107	plasmidom	680	15	57.7	122	2	Q692D8_CVHBA	Q692D8	sars corona
608	15	57.7	98	2	Q4UZT6_XANCP	Q4UZT6	xanthomonas	681	15	57.7	122	2	Q6GYO5_CVHBA	Q6GYO5	sars corona
609	15	57.7	98	2	Q8PDJ5_XANCP	Q8PDJ5	xanthomonas	682	15	57.7	122	2	Q6R7Y0_CVHBA	Q6R7Y0	sars corona
610	15	57.7	98	2	Q90D13_9H1V1	Q90D13	human immun	683	15	57.7	122	2	Q6S8D1_CVHBA	Q6S8D1	sars corona
611	15	57.7	99	2	Q5YAU4_AGBOR	Q5YAU4	agelena ori	684	15	57.7	122	2	Q6S288_CVHSA	Q6S288	sars corona
612	15	57.7	99	2	Q71267_9H1V1	Q71267	human immun	685	15	57.7	122	2	Q6R1D6_CVHSA	Q6R1D6	sars corona
613	15	57.7	99	2	Q90D11_9H1V1	Q90D11	human immun	686	15	57.7	122	2	Q76XZ9_CVHSA	Q76XZ9	sars corona
614	15	57.7	99	2	Q90D12_9H1V1	Q90D12	human immun	687	15	57.7	122	2	Q6RCV9_CVHSA	Q6RCV9	sars corona
615	15	57.7	100	2	Q90D15_9H1V1	Q90D15	human immun	688	15	57.7	122	2	Q6RCX0_CVHSA	Q6RCX0	sars corona

689	15	57.7	122	2	06RCV1 CVHSA	06rcv1	sars	corona	762	15	57.7	136	2	04V6K1 DROME	04v6k1	drosophila
690	15	57.7	122	2	06RCZ2 CVHSA	06rcz2	sars	corona	763	15	57.7	136	2	06GRP1 DROVA	06grp1	oryza sativ
691	15	57.7	122	2	06RD03 CVHSA	06rd03	sars	corona	764	15	57.7	137	2	06XUB8 DROVA	06xub8	drosophila
692	15	57.7	122	2	06RD14 CVHSA	06rd14	sars	corona	765	15	57.7	137	2	063U00 BURPS	063u00	burkholderi
693	15	57.7	122	2	06RD25 CVHSA	06rd25	sars	corona	766	15	57.7	137	2	05FXE1 HEHPC	05fxe1	hepatitis c
694	15	57.7	122	2	06RD36 CVHSA	06rd36	sars	corona	767	15	57.7	137	2	05FXE2 HEHPC	05fxe2	hepatitis c
695	15	57.7	122	2	06RD47 CVHSA	06rd47	sars	corona	768	15	57.7	137	2	05FXF2 HEHPC	05fxf2	hepatitis c
696	15	57.7	122	2	06RD58 CVHSA	06rd58	sars	corona	769	15	57.7	137	2	05FXF8 HEHPC	05fxf8	hepatitis c
697	15	57.7	122	2	06SR05 CVHSA	06sr05	sars	corona	770	15	57.7	137	2	05FXG5 HEHPC	05fxg5	hepatitis c
698	15	57.7	122	2	06SR06 CVHSA	06sr06	sars	corona	771	15	57.7	137	2	05FXG9 HEHPC	05fxg9	hepatitis c
699	15	57.7	122	2	06SR05 CVHSA	06sr05	sars	corona	772	15	57.7	137	2	05FXH0 HEHPC	05fxh0	hepatitis c
700	15	57.7	122	2	06SR05 CVHSA	06sr05	sars	corona	773	15	57.7	137	2	05FXH1 HEHPC	05fxh1	hepatitis c
701	15	57.7	122	2	06SR05 CVHSA	06sr05	sars	corona	774	15	57.7	137	2	05FXH2 HEHPC	05fxh2	hepatitis c
702	15	57.7	122	2	06SR05 CVHSA	06sr05	sars	corona	775	15	57.7	137	2	05FXH3 HEHPC	05fxh3	hepatitis c
703	15	57.7	122	2	06SR05 CVHSA	06sr05	sars	corona	776	15	57.7	137	2	05FXH4 HEHPC	05fxh4	hepatitis c
704	15	57.7	122	2	06SR05 CVHSA	06sr05	sars	corona	777	15	57.7	137	2	05FXH5 HEHPC	05fxh5	hepatitis c
705	15	57.7	122	2	06SR05 CVHSA	06sr05	sars	corona	778	15	57.7	137	2	05FXH6 HEHPC	05fxh6	hepatitis c
706	15	57.7	122	2	06SRV0 CVHSA	06srvo	sars	corona	779	15	57.7	137	2	05FXH7 HEHPC	05fxh7	hepatitis c
707	15	57.7	122	2	06SRV5 CVHSA	06srvs	sars	corona	780	15	57.7	137	2	04S169 TETNG	04s169	tetradon n
708	15	57.7	122	2	06SRW0 CVHSA	06srw0	sars	corona	781	15	57.7	138	2	06Y9V8 LYCRS	06y9v8	lycopersico
709	15	57.7	122	2	06SRW5 CVHSA	06srw5	sars	corona	782	15	57.7	138	2	07FEB7 RACME	07feb7	ribes cyto
710	15	57.7	122	2	06SRX0 CVHSA	06srx0	sars	corona	783	15	57.7	139	2	06GY53 XANOR	06gy53	xanthomonas
711	15	57.7	122	2	06SRX5 CVHSA	06srx5	sars	corona	784	15	57.7	139	2	04MNG4 BACCB	04mng4	bacillus ce
712	15	57.7	122	2	06SRV0 CVHSA	06srvo	sars	corona	785	15	57.7	139	2	06HFJ1 BACBK	06hfj1	bacillus th
713	15	57.7	122	2	06SRV5 CVHSA	06srvs	sars	corona	786	15	57.7	139	2	0733N2 BACG1	0733n2	bacillus ce
714	15	57.7	122	2	06SRZ0 CVHSA	06srz0	sars	corona	787	15	57.7	139	2	081AG4 BACCR	081ag4	bacillus ce
715	15	57.7	122	2	06VA72 CVHSA	06va72	sars	corona	788	15	57.7	139	2	097D28 CLOAB	097d28	clostridium
716	15	57.7	122	2	06VA83 CVHSA	06va83	sars	corona	789	15	57.7	139	2	081Y90 BACAN	081y90	bacillus an
717	15	57.7	122	2	06VA94 CVHSA	06va94	sars	corona	790	15	57.7	139	2	0637J9 BACZ2	0637j9	bacillus ce
718	15	57.7	122	2	04JDP8 CVHSA	04jdp8	sars	corona	791	15	57.7	139	2	069963 9H1V1	069963	human immun
719	15	57.7	122	2	04JDR1 CVHSA	04jdr1	sars	corona	792	15	57.7	140	1	LY6H HUMAN	ly6h	human
720	15	57.7	122	2	06GJ49 XENLA	06gj49	xenopus lae		793	15	57.7	140	2	06AXX0 HUMAN	06axx0	homo sapien
721	15	57.7	123	2	09JQ22 NEIMA	09jq22	neisseria m		794	15	57.7	140	2	04V3U1 DROME	04v3u1	drosophila
722	15	57.7	123	2	07DDH5 NEIMB	07ddh5	neisseria m		795	15	57.7	140	2	06GZ25 ORYSA	06gz25	oryza sativ
723	15	57.7	123	2	06Y908 9H1V1	06y908	human immun		796	15	57.7	140	2	08AR11 9H1V1	08ar11	human immun
724	15	57.7	124	2	09PKS3 CHLMU	09pks3	chlamydia m		797	15	57.7	141	2	06GKZ6 DROME	06gkz6	drosophila
725	15	57.7	124	2	07SR41 9H1V1	07sr41	human immun		798	15	57.7	141	2	0664U8 YERPS	0664u8	yersinia ps
726	15	57.7	125	2	06KDP2 ECOLI	06kdp2	escherichia		799	15	57.7	141	2	08ZJ85 YERPE	08zj85	yersinia pe
727	15	57.7	125	2	04KFP0 PSBFS	04kfp0	psuedomonas		800	15	57.7	142	2	09YF15 DROME	09yf15	drosophila
728	15	57.7	125	2	08F1Z3 ECOL6	08f1z3	escherichia		801	15	57.7	143	2	06ZTJ0 HUMAN	06ztj0	homo sapien
729	15	57.7	125	2	04TEY5 TETNG	04tey5	tetradon n		802	15	57.7	143	2	08A406 HORVD	08a406	hoddeum vul
730	15	57.7	125	2	06RUW8 9H1V1	06ruw8	human immun		803	15	57.7	143	2	09ACU2 DROME	09acu2	strepomyce
731	15	57.7	126	2	04ITX6 GIBZE	04itx6	gibberella		804	15	57.7	143	2	06MH33 BDEBA	06mh33	babelliovibri
732	15	57.7	126	2	07R635 GIALLA	07r635	giardia lam		805	15	57.7	144	2	04F8J1 EUCGL	04f8j1	eucalyptus
733	15	57.7	126	2	088C21 PSBPK	088c21	pseudomonas		806	15	57.7	144	2	08CZ18 YERPE	08cz18	yersinia pe
734	15	57.7	127	2	04ZL69 PSBSY	04zl69	pseudomonas		807	15	57.7	145	2	08WQ22 LOCME	08wq22	locusta mg
735	15	57.7	127	2	069962 9H1V1	069962	human immun		808	15	57.7	145	2	09VZT5 DROME	09vzt5	drosophila
736	15	57.7	128	1	KRA24 HUMAN	08lbo3	arabid		809	15	57.7	145	2	06K6Z6 ORYSA	06k6z6	oryza sativ
737	15	57.7	128	2	08LBO3 ARATH	08lbo3	arabidopsis		810	15	57.7	146	2	05K4F7 SCGJR	05k4f7	schistoscer
738	15	57.7	128	2	04ZU08 PSBSY	04zu08	pseudomonas		811	15	57.7	146	2	07XNZ5 ORYSA	07xnz5	oryza sativ
739	15	57.7	129	2	05D9M3 SCHJA	05d9m3	schistosoma		812	15	57.7	146	2	08AXW7 MICCO	08axw7	micritus co
740	15	57.7	129	2	06IKR4 DROME	06ikr4	drosophila		813	15	57.7	146	2	04RFX5 TETNG	04rfx5	tetradon n
741	15	57.7	129	2	05OK95 LYNST	05ok95	lynaea ste		814	15	57.7	147	2	05Y2U3 9H1V1	05y2u3	human immun
742	15	57.7	129	2	04IVL5 AZOVI	04ivl5	azocobacter		815	15	57.7	148	2	06ZMW2 HUMAN	06zwm2	homo sapien
743	15	57.7	130	2	054OR7 DICDI	054or7	dicyosteli		816	15	57.7	148	2	05CTC71 HUMAN	05ctc71	homo sapien
744	15	57.7	130	2	06THB0 DROME	06thb0	drosophila		817	15	57.7	148	2	04RLJ3 TETNG	04rlj3	tetradon n
745	15	57.7	131	2	09S237 STRCO	09s237	strepomyce		818	15	57.7	149	2	08UAT1 AGRT5	08uat1	agrobacteri
746	15	57.7	131	2	06CZ68 ERWCT	06cz68	erythra car		819	15	57.7	149	2	0898V0 CIOTE	0898v0	ciostriidum
747	15	57.7	132	2	05SZH5 CRYNCE	05szh5	cryptococcu		820	15	57.7	150	2	05ACZ8 CANAL	05acz8	candida alb
748	15	57.7	132	2	05D995 SCHJA	05d995	schistosoma		821	15	57.7	150	2	06ICF4 CAENHAB1	06icf4	caenorhabdi
749	15	57.7	132	2	09AN59 BRABA	09an59	bradyrhizob		822	15	57.7	150	2	06IDZ5 PARTE	06idz5	paramacium
750	15	57.7	132	2	07SOV8 9H1V1	07sov8	human immun		823	15	57.7	151	2	09BHM5 CABEL	09bhm5	caenorhabdi
751	15	57.7	133	2	04TZU4 TETNG	04tzu4	tetradon n		824	15	57.7	152	2	08TIR4 METNAC	08tir4	methanosarc
752	15	57.7	133	2	08U0U5 9H1V1	08u0u5	human immun		825	15	57.7	152	2	07PXZ2 ANOGA	07pxz2	anopheles g
753	15	57.7	134	2	024959 GIALLA	024959	giardia lam		826	15	57.7	153	2	080876 BSMV	080876	barley etci
754	15	57.7	134	2	024988 GIALLA	024988	giardia lam		827	15	57.7	153	2	0941T8 ORYSA	0941t8	oryza sativ
755	15	57.7	134	2	04V6M8 DROME	04v6m8	drosophila		828	15	57.7	154	2	05TPK5 ANOGA	05tpk5	anopheles g
756	15	57.7	134	2	06RUX2 9H1V1	06rux2	human immun		829	15	57.7	155	2	09SKW0 BOVIN	09skw0	boe taurus
757	15	57.7	135	1	YDQA SCHPO	014J02	schizosacch		830	15	57.7	155	2	08GSR9 WHEAT	08gsr9	wheat
758	15	57.7	135	2	05D8E6 SCHJA	05d8e6	schistosoma		831	15	57.7	155	2	069536 9BETA	069536	human herpes
759	15	57.7	135	2	04QIR1 LEJIMA	04qir1	leishmania		832	15	57.7	155	2	01Z317 9H1V1	01z317	human immun
760	15	57.7	135	2	084YV6 ORYSA	084yv6	oryza sativ		833	15	57.7	156	2	08QW84 9NUCL	08qw84	manebstra co
761	15	57.7	135	2	08H6Y5 PHYTIN	08h6y5	phytophthor		834	15	57.7	157	1	CI0255 HUMAN	ci0255	homo sapien

835	15	57.7	157	1	C1025_MOUSE	Q9d772	mus musculus	908	15	57.7	178	2	Q66M85_9HEPC	Q66M85	hepaticis c
836	15	57.7	157	2	Q51W61_MAGOR	Q51W61	magnaporthie	909	15	57.7	179	2	Q9XU28_CAEEL	Q9XU28	caenorhabdi
837	15	57.7	157	2	Q8BNR9_MOUSE	Q8BNR9	mus musculus	910	15	57.7	179	2	Q82SM7_NITEU	Q82SM7	nitrosomona
838	15	57.7	158	2	Q9R8C6_YERPE	Q9R8C6	yerinia pe	911	15	57.7	180	2	Q4LRE5_9BURX	Q4LRE5	burholderi
839	15	57.7	158	2	Q74Y15_YERPE	Q74Y15	yerinia pe	912	15	57.7	180	2	Q9U2M4_CAEEL	Q9U2M4	caenorhabdi
840	15	57.7	158	2	Q8C4G8_MOUSE	Q8C4G8	mus musculus	913	15	57.7	181	2	Q66ME0_9HEPC	Q66ME0	hepaticis c
841	15	57.7	158	2	Q5XQV5_9HEPC	Q5XQV5	hepaticis c	914	15	57.7	182	1	CX036_HUMAN	Q9h710	homo sapien
842	15	57.7	158	2	Q5XQ21_9HEPC	Q5XQ21	hepaticis c	915	15	57.7	182	2	Q9CR91_MOUSE	Q9CR91	mus muscu
843	15	57.7	158	2	Q4S2N4_TENG	Q4S2N4	tetradon n	916	15	57.7	183	2	Q5T595_HUMAN	Q5T595	giardia lam
844	15	57.7	159	2	Q8BXD9_MOUSE	Q8BXD9	mus musculus	917	15	57.7	183	2	Q7ON08_GILTA	Q7ON08	burholderi
845	15	57.7	159	2	Q58UJ7_9HEPC	Q58UJ7	hepaticis c	918	15	57.7	183	2	Q4LGV4_9BURX	Q4LGV4	burholderi
846	15	57.7	159	2	Q58UJ8_9HEPC	Q58UJ8	hepaticis c	919	15	57.7	184	1	IMF3_MOUSE	Q92LY2	mus musculus
847	15	57.7	159	2	Q58UJ9_9HEPC	Q58UJ9	hepaticis c	920	15	57.7	184	2	Q57Q73_SALCH	Q57Q73	salmonella
848	15	57.7	160	2	Q7CSF5_AGRYS	Q7CSF5	agrobacteri	921	15	57.7	185	2	Q7RUQ3_NEUCR	Q7RUQ3	nitrosospora
849	15	57.7	160	2	Q98AU9_RHIL0	Q98AU9	rhizobium l	922	15	57.7	185	2	Q5N207_AZOSR	Q5N207	azarcus sp
850	15	57.7	161	2	Q4R5M8_MACPA	Q4R5M8	macaca fasc	923	15	57.7	185	2	Q9D6L6_MOUSE	Q9D6L6	mus musculus
851	15	57.7	162	2	Q5Z408_ORYSA	Q5Z408	oryza sativ	924	15	57.7	185	2	Q7ZG22_9HIV1	Q7ZG22	human immun
852	15	57.7	163	2	Q4S851_CAEEL	Q4S851	caenorhabdi	925	15	57.7	186	1	C1025_BRABE	Q7ZG22	human immun
853	15	57.7	163	2	Q5YFN8_VIRIU	Q5YFN8	singapore g	926	15	57.7	186	2	Q4IE25_GIBZE	Q4IE25	gibberellia
854	15	57.7	164	2	Q8N8L1_HUMAN	Q8N8L1	homo sapien	927	15	57.7	186	2	Q6ZTD4_HUMAN	Q6ZTD4	homo sapien
855	15	57.7	164	2	Q9DC82_MOUSE	Q9DC82	mus musculus	928	15	57.7	186	2	Q5N717_ORYSA	Q5N717	oryza sativ
856	15	57.7	165	2	Q5I6Z4_CHIAB	Q5I6Z4	chlamydomphi	929	15	57.7	186	2	Q7ZG29_9HIV1	Q7ZG29	human immun
857	15	57.7	166	1	ZCH13_HUMAN	ZCH13	homo sapien	930	15	57.7	187	2	Q5TSD5_ANOGA	Q5TSD5	anopheles g
858	15	57.7	166	2	Q5T594_HUMAN	Q5T594	homo sapien	931	15	57.7	188	2	Q6ESN8_ORYSA	Q6ESN8	oryza sativ
859	15	57.7	166	2	Q6Z4B6_ORYSA	Q6Z4B6	oryza sativ	932	15	57.7	188	2	Q92XN0_RHIME	Q92XN0	rhizobium m
860	15	57.7	166	2	Q8BPJ3_MOUSE	Q8BPJ3	mus musculus	933	15	57.7	189	2	Q7ACX1_ECO57	Q7ACX1	escherichia
861	15	57.7	166	2	Q80657_9HIV1	Q80657	human immun	934	15	57.7	189	2	Q4SMJ5_TETNG	Q4SMJ5	tetradon n
862	15	57.7	167	2	Q52D68_MAGGR	Q52D68	magnaporthie	935	15	57.7	189	2	Q7ZG13_9HIV1	Q7ZG13	human immun
863	15	57.7	167	2	Q6YRN6_ORYSA	Q6YRN6	oryza sativ	936	15	57.7	190	2	Q6TH54_DROME	Q6TH54	droscophila
864	15	57.7	167	2	Q89YU6_BACTN	Q89YU6	bacteroides	937	15	57.7	190	2	Q67WM2_ORYSA	Q67WM2	oryza sativ
865	15	57.7	168	1	RUVIC_CHICV	RUVIC	chlamydomphi	938	15	57.7	190	2	Q80WK2_ICTPU	Q80WK2	ictalurus p
866	15	57.7	168	2	Q8IOB0_DROME	Q8IOB0	droscophila	939	15	57.7	191	2	P87Z66_YEAST	P87Z66	saccharomyc
867	15	57.7	168	2	Q6IH00_MOUSE	Q6IH00	mus musculus	940	15	57.7	191	2	Q5A799_CANAL	Q5A799	canidia alb
868	15	57.7	168	2	Q8SD21_9CAUD	Q8SD21	pseudomonas	941	15	57.7	191	2	Q502H7_BRABE	Q502H7	brachydanio
869	15	57.7	168	2	Q7XK29_ORYSA	Q7XK29	oryza sativ	942	15	57.7	191	2	Q39320_9HIV1	Q39320	human immun
870	15	57.7	168	2	Q12320_9HIV1	Q12320	homo sapien	943	15	57.7	192	2	P91715_DUGIT	P91715	dugesiopsis t
871	15	57.7	169	1	KRA59_HUMAN	P26371	homo sapien	944	15	57.7	192	2	Q9LPI6_ARATH	Q9LPI6	arabidopsis
872	15	57.7	169	2	Q6H9I5_9ARCH	Q6H9I5	uncultured	945	15	57.7	193	2	Q9N1D8_MACMA	Q9N1D8	macaca mula
873	15	57.7	169	2	Q5SZU8_HUMAN	Q5SZU8	choristoneu	946	15	57.7	193	2	Q7UXS2_RHOBA	Q7UXS2	rhodopitrell
874	15	57.7	169	2	Q12331_9HIV1	Q12331	human immun	947	15	57.7	194	2	Q89KY3_BRADY	Q89KY3	bradyrhizob
875	15	57.7	170	2	Q9XXZ9_DUGOA	Q9XXZ9	dugesia jap	948	15	57.7	194	2	Q5K266_GUTTH	Q5K266	gulliardia
876	15	57.7	170	2	Q5GQK9_MOUSE	Q5GQK9	mus musculus	949	15	57.7	195	2	Q17630_CAEEL	Q17630	caenorhabdi
877	15	57.7	170	2	Q910I3_HCMW	Q910I3	human cytom	950	15	57.7	195	2	Q7SHK2_ORYSA	Q7SHK2	oryza sativ
878	15	57.7	170	2	Q6VTV8_NPCVD	Q6VTV8	choristoneu	951	15	57.7	195	2	Q68291_BORAD	Q68291	borrelia an
879	15	57.7	170	2	Q12316_9HIV1	Q12316	human immun	952	15	57.7	195	2	Q92K71_RHIME	Q92K71	rhizobium m
880	15	57.7	170	2	Q12321_9HIV1	Q12321	human immun	953	15	57.7	195	2	Q9GX82_BR	Q9GX82	rattus norv
881	15	57.7	170	2	Q12326_9HIV1	Q12326	human immun	954	15	57.7	196	2	Q8T3B0_CAEEL	Q8T3B0	caenorhabdi
882	15	57.7	170	2	Q12330_9HIV1	Q12330	human immun	955	15	57.7	196	2	Q4QKX7_HAEI8	Q4QKX7	haemophilus
883	15	57.7	171	1	Y041_NPVOF	Q10301	orygia pseu	956	15	57.7	196	2	Q4RVM0_TETNG	Q4RVM0	tetradon n
884	15	57.7	171	2	Q68H17_NPVAP	Q68H17	antherea p	957	15	57.7	198	2	Q51I19_MACGR	Q51I19	magnaporthie
885	15	57.7	171	2	Q12314_9HIV1	Q12314	human immun	958	15	57.7	198	2	Q4SAD0_TETNG	Q4SAD0	tetradon n
886	15	57.7	171	2	Q12315_9HIV1	Q12315	human immun	959	15	57.7	199	2	Q6YTV8_ORYSA	Q6YTV8	oryza sativ
887	15	57.7	171	2	Q12319_9HIV1	Q12319	human immun	960	15	57.7	199	2	Q8X2A8_ECO57	Q8X2A8	ecocherichia
888	15	57.7	171	2	Q12322_9HIV1	Q12322	human immun	961	15	57.7	199	2	Q8ZNI0_SALTY	Q8ZNI0	salmonella
889	15	57.7	171	2	Q12323_9HIV1	Q12323	human immun	962	15	57.7	199	2	Q5WSF7_LEBPL	Q5WSF7	legionella
890	15	57.7	171	2	Q12324_9HIV1	Q12324	human immun	963	15	57.7	199	2	Q5XON2_LEBPA	Q5XON2	legionella
891	15	57.7	171	2	Q12325_9HIV1	Q12325	human immun	964	15	57.7	199	2	Q5ZCR90_LEBPA	Q5ZCR90	legionella
892	15	57.7	171	2	Q12327_9HIV1	Q12327	human immun	965	15	57.7	199	2	Q8X3M1_ECO57	Q8X3M1	ecocherichia
893	15	57.7	172	2	Q8W287_HUMAN	Q8W287	homo sapien	966	15	57.7	199	2	Q9QNZ5_9HIV1	Q9QNZ5	human immun
894	15	57.7	172	2	Q5T590_HUMAN	Q5T590	homo sapien	967	15	57.7	200	1	RECR_CHILW	Q9PFI4	chlamydia m
895	15	57.7	172	2	Q5UB16_PHYCI	Q5UB16	phytophthor	968	15	57.7	200	2	Q84UR5_ARALY	Q84UR5	arabidopsis
896	15	57.7	174	2	Q8TRD7_TASBA	Q8TRD7	taenia sagi	969	15	57.7	201	2	Q6FPU3_GLUOX	Q6FPU3	gluconobact
897	15	57.7	174	2	Q8CK86_YERPE	Q8CK86	yerinia pe	970	15	57.7	201	2	Q6ZNU3_HUMAN	Q6ZNU3	homo sapien
898	15	57.7	174	2	Q8BURS_MOUSE	Q8BURS	mus musculus	971	15	57.7	201	2	Q8G669_BIRLO	Q8G669	bifidobacte
899	15	57.7	175	2	Q6IWT5_PAPPR	Q6IWT5	paramectum	972	15	57.7	202	2	Q5A2D9_EMBETI	Q5A2D9	aspergillus
900	15	57.7	175	2	Q57ZJ5_PHYIN	Q57ZJ5	phytophthor	973	15	57.7	202	2	Q7AFI2_GEOSL	Q7AFI2	geobacteri
901	15	57.7	176	2	Q8XQB3_RALSO	Q8XQB3	ralstonia s	974	15	57.7	203	2	Q8N8X7_HUMAN	Q8N8X7	homo sapien
902	15	57.7	176	2	Q9IGL3_NPYEP	Q9IGL3	epiphyas po	975	15	57.7	203	2	Q61G66_CAEBR	Q61G66	caenorhabdi
903	15	57.7	177	1	RELX_MESAU	Q64I71	mesocricetu	976	15	57.7	203	2	Q84TB9_ORYSA	Q84TB9	oryza sativ
904	15	57.7	177	2	Q363I3_9HIV1	Q363I3	human immun	977	15	57.7	204	2	Q5CJ96_CRYHO	Q5CJ96	cryptospori
905	15	57.7	178	2	Q16423_CAEEL	Q16423	caenorhabdi	978	15	57.7	204	2	Q5JM57_ORYSA	Q5JM57	oryza sativ
906	15	57.7	178	2	Q4QYX6_HAECO	Q4QYX6	haemonchus	979	15	57.7	205	2	Q4IWM5_PLABE	Q4IWM5	plasmodium
907	15	57.7	178	2	Q6P064_MOUSE	Q6P064	mus musculus	980	15	57.7	205	2	Q6ZBC3_ORYSA	Q6ZBC3	oryza sativ

```
961 15 57.7 207 2 Q4VE68 DROME
962 15 57.7 207 2 Q70016 SHIDY
963 15 57.7 207 2 Q8FCJ3 ECOL6
964 15 57.7 207 2 Q8XSN4 ECOS7
965 15 57.7 208 2 Q6XHV9 DROVA
966 15 57.7 208 2 Q82069 SOLTU
967 15 57.7 208 2 Q4F8J3 EUGCL
968 15 57.7 208 2 Q9AHJ9 YLACO
969 15 57.7 209 2 Q55GM2 DICI1
990 15 57.7 210 2 Q7KML7 DICI1
991 15 57.7 210 2 Q7XU70 ORYSA
992 15 57.7 210 2 Q93CB9 MYCBA
993 15 57.7 210 2 Q39318 SHIV1
994 15 57.7 211 1 G12 ANOGA
995 15 57.7 211 2 Q7Q5V5 ANOGA
996 15 57.7 211 2 Q50SC9 ENTHI
997 15 57.7 211 2 Q6WES3 9TELE
998 15 57.7 212 2 Q50PE5 ENTHI
999 15 57.7 212 2 Q9XV06 CAEEL
1000 15 57.7 212 2 Q4Q6X5 LETMA
Q4VE68 drosophila
Q70016 shigella cy
Q8FCJ3 escherichia
Q8XSN4 escherichia
Q6XHV9 drosophila
Q82069 solanum tub
Q4F8J3 eucalyptus
Q9AHJ9 laccobacill
Q55GM2 dictyostel
Q7KML7 dictyostel
Q7XU70 oryza sativ
Q93CB9 mycobacteri
Q39318 human immun
Q17040 anopheles g
Q7Q5V5 anopheles g
Q50SC9 entamoeba h
Q6WES3 polliachus
Q50PE5 entamoeba h
Q9XV06 caenorhabd
Q4Q6X5 leishmania
```

## ALIGNMENTS

```
RESULT 1
ID Q96KM3 HUMAN PRELIMINARY; PRT; 213 AA.
```

```
AC Q96KM3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Aberrant WW domain-containing oxidoreductase.
GN Name=WWOX;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCB1_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21457335; PubMed=11572989; DOI=10.1073/pnas.191175896;
RA Palge A.J.W., Taylor K.J., Taylor C., Hillier S.G., Farrington S.,
RT Scott D., Porteous D.J., Smyth J.F., Gabra H., Watson J.E.V.,
RT "WWOX: a candidate tumor suppressor gene involved in multiple tumor
RT types.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:11417-11422(2001).
RL EMBL; AF325432; AAL05451.1; -; Genomic DNA.
DR EMBL; AF325423; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325424; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325426; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325433; AAL05451.1; JOINED; Genomic DNA.
DR EMBL; AF325425; AAL05451.1; JOINED; Genomic DNA.
DR HSSP; Q13526; IPTN.
DR Ensembl; ENSG00000186153; Homo sapiens.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS01159; WW_DOMAIN 1; 2.
DR PROSITE; PSS0020; WW_DOMAIN 2; 2.
SQ SEQUENCE 213 AA; 23866 MW; A21054FF8214CC7C CRC64;
```

```
Query Match 69.2%; Score 18; DB 2; Length 213;
Best Local Similarity 20.0%; Pred. No. 0.43;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 CXXXXXXXXXC 10
DB 184 CTSTTAAAC 193
```

```
RESULT 2
Q586F4_9TRYP
```

```
ID Q586F4_9TRYP PRELIMINARY; PRT; 263 AA.
```

```
AC Q586F4;
DT 10-MAY-2005 (TREMBLrel. 30, Created)
DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=TB927.6.4820;
OS Trypanosoma brucei.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCB1_TaxID=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUtatio.1;
RA Chedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
RA Shalton J., Hou L., Djikeng A., Feldblum T., Hostetler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pal G., Peterson J.,
RA Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wanstless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUtatio.1;
RA El-Sayed N.M., Khalak H., Adams M.D.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUtatio.1;
RA Haas B., Blandin G., El-Sayed N.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC008146; AAX80300.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 263 AA; 28873 MW; F61DF5AF910531AB CRC64;
```

```
Query Match 69.2%; Score 18; DB 2; Length 263;
Best Local Similarity 20.0%; Pred. No. 0.46;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 CXXXXXXXXXC 10
DB 95 CATTSSSSAC 104
```

## RESULT 3

```
ID Q7PRO7 ANOGA PRELIMINARY; PRT; 265 AA.
```

```
AC Q7PRO7;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE ENSANGP0000001657 (Fragment).
GN ORFNames=ENSANGG0000001387;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anopheles; Anopheles.
OX NCB1_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008847; EAA06779.3; -; Genomic DNA.
FT NON_TER 1 1
FT 269 269
```

SEQ SEQUENCE 269 AA; 24228 MW; 25BBPF71FD71F1F2 CRC64;  
Query Match 69.2%; Score 18; DB 2; Length 269;  
Best Local Similarity 20.0%; Pred. No. 0.47;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXC 10  
DB 149 CTSSTTSC 158

RESULT 4  
OSTX10 ANOCA PRELIMINARY; PRT; 281 AA.  
ID OSTX10;  
AC 01-FEB-2005 (TREMBlrel. 29, Created)  
DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)  
DE ENSANGP0000027659 (Fragment).  
GN ORFNames=ENSANG0000007023;  
OS Anopheles gambiae str. PEST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;  
OC Anophelinae; Anopheles.  
OX NCBI\_TaxId=180454;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=PEST;  
RT "The Anopheles gambiae Sequence Committee;  
"Anopheles gambiae re-annotation."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=PEST;  
RT The Anopheles gambiae Sequence Committee;  
Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
CC -! CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AAB01008807; EAL41760.1; -; Genomic\_DNA.  
FT NON\_TER 1  
FT TER 281  
SQ SEQUENCE 281 AA; 25341 MW; 0462E1169FB642B CRC64;

Query Match 69.2%; Score 18; DB 2; Length 281;  
Best Local Similarity 20.0%; Pred. No. 0.48;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXC 10  
DB 107 CTRAASTAC 116

RESULT 5  
YMS8 YEAST STANDARD; PRT; 313 AA.  
ID YMS8 YEAST  
AC 003655;  
DT 01-NOV-1997 (rel. 35, Created)  
DT 01-NOV-1997 (rel. 35, Last sequence update)  
DT 10-MAY-2005 (rel. 47, Last annotation update)  
DE Hypothetical 35.0 kDa protein in PFK2-HFA1 intergenic region.  
GN OrderedAccession=YMR206W; ORFNames=YMR325.07;  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;  
OC NCBI\_TaxId=4932;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=S288C/ AB972;  
RA MEDLINE=9713268; PubMed=9169872;  
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,  
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,  
RA Jags K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,

RA Rice P., Skelton J., Walsh S.V., Whitehead S., Barrell B.G.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome  
XIII."  
RL Nature 387:90-93 (1997).  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC EMBL; 248755; CAA8648.1; -; Genomic\_DNA.  
DR PIR; S59448; S59448.  
DR GenBank; 142881; -;  
DR Ensembl; YMR206W; Saccharomyces cerevisiae.  
DR SGD; S000004819; YMR206W.  
KM Complete proteome; Hypothetical protein.  
FT COMPBIAS 3  
FT COMPBIAS 146 149 Poly-Ser.  
FT COMPBIAS 246 252 Poly-Gln.  
FT COMPBIAS 246 252 Poly-Ser.  
SQ SEQUENCE 313 AA; 35018 MW; 9D92BFD8982577F0 CRC64;

Query Match 69.2%; Score 18; DB 1; Length 313;  
Best Local Similarity 20.0%; Pred. No. 0.5;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXC 10  
DB 245 CTSSTTSC 254

RESULT 6  
O6ZDR4 ORYSA PRELIMINARY; PRT; 357 AA.  
ID O6ZDR4 ORYSA  
AC O6ZDR4;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Putative transcription factor Myb protein.  
GN Name=P0481F05.20;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxId=39947;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, PAC  
clone: P0481F05."  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
CC -! SUBCELLULAR LOCATION: Nuclear (By similarity).  
DR EMBL; AP004376; BAD09322.1; -; Genomic\_DNA.  
DR HSSP; Q03237; IASU.  
DR Gramene; O6ZDR4; -;  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR InterPro; IPR012287; Homeodomain-rel.  
DR InterPro; IPR001005; MYB\_DNA\_bd.  
DR Pfam; PF00249; MYB\_DNA-binding; 2.  
DR SMART; SM00717; SANT; 2.  
DR PROSITE; PS00037; MYB\_1; UNKNOWN\_1.  
DR PROSITE; PS00034; MYB\_2; 1.  
DR PROSITE; PS00090; MYB\_3; 2.  
KW Nuclear protein; Repeat.  
SQ SEQUENCE 357 AA; 36754 MW; 3FAF56860D3BAE22 CRC64;

Query Match 69.2%; Score 18; DB 2; Length 357;  
Best Local Similarity 20.0%; Pred. No. 0.52;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXC 10



DB 322 CSATASASAC 331

RESULT 7

ID 07010 ANOCA PRELIMINARY; PRT; 362 AA.

AC 070101

DT 01-MAR-2004 (TREMBlrel. 26, Created)

DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE ENSANGP0000009383 (Fragment).

GN ORFNames=ENSANGG0000007023;

OS Anopheles gambiae str. PEST.

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;

OC Anophelinae; Anopheles.

OX NCBI\_TaxID=180454;

RM [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RT "Anopheles gambiae re-annotation."

RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.

RM [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=PEST;

RG The Anopheles gambiae Sequence Committee;

RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.

CC -! CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; AAAB01008807; EAA04378.2; -; Genomic\_DNA.

FT NON\_TER 1

SQ SEQUENCE 362 AA; 32490 MM; 384376DDFA31BEC0 CRC64;

Query Match 69.2%; Score 18; DB 2; Length 362;

Best Local Similarity 20.0%; Pred. No. 0.52;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10

DB 187 CTAASSTAC 196

RESULT 8

Q9RV87 DEIRA PRELIMINARY; PRT; 480 AA.

ID Q9RV87 DEIRA PRELIMINARY; PRT; 480 AA.

AC Q9RV87

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Hypothetical protein DR1142.

GN OrderedlocusNames=DR1142.

OS Deinococcus radiodurans.

OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;

OC Deinococcaceae; Deinococcus.

OX NCBI\_TaxID=1299;

RM [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=1 / ATCC 13939 / DSM 20539 / NCIB 9279;

RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.U., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,

RA Vamathevan J.U., Lam P., McDonald L.A., Ueberback T.R., Zalewski C.,

RA Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,

RA Fraser C.W.;

RT "Genome sequence of the radioresistant bacterium Deinococcus

RT radiodurans R1."

RL Science 286:1571-1577(1999).

DR EMBL; AE001963; AAP10716.1; -; Genomic\_DNA.

PIR; E75433; E75433.

DR TIGR: DR1142; -

KM Complete proteome; Hypothetical protein.

SQ SEQUENCE 480 AA; 49364 MM; 189624DAAC24AB CRC64;

Query Match 69.2%; Score 18; DB 2; Length 480;

Best Local Similarity 20.0%; Pred. No. 0.58;

Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXX 10

DB 464 CASAAATAC 473

RESULT 9

ID TWCC1 HUMAN STANDARD; PRT; 653 AA.

AC 094876; Q68E06; Q81XM8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 10-MAY-2005 (Rel. 47, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Transmembrane and coiled-coil domains protein 1.

GN Name=TWCC1; Synonyms=K1A0779;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;

OC Homo

OX NCBI\_TaxID=9606;

RM [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RC TISSUE=Skin;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Rana S.S., Loquellano N.J., Peter G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McSwan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Gardia A.M., Gay L.J., Hultk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RM [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 15-653.

RC TISSUE=Endometrial tumor;

RG The German CDNA consortium;

RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.

RM [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 334-653.

RC TISSUE=Brain;

RX MEDLINE=99087487; PubMed=9872452;

RA Nagae T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,

RA Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. XI.

RT The complete sequences of 100 new cDNA clones from brain which code

RT for large proteins in vitro."

RL DNA Res. 5:277-286(1998).

CC -! SIMILARITY: Belongs to the TEX28 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not

CC removed.



```

CC -----
DR EMBL; AB018322; BAA3449.1; -; mRNA.
DR EMBL; BC039859; AAH39859.1; -; mRNA.
DR EMBL; CR749206; CAH18064.2; -; mRNA.
DR Ensemble; ENSG00000172765; Homo sapiens.
DR HGNC; HGNC:29116; TMCC1.
KW Coiled coil; Transmembrane.
FT TRANSMEM 592 612 Potential.
FT TRANSMEM 625 645 Potential.
FT COILED 228 313 Potential.
FT COILED 458 576 Potential.
SQ SEQUENCE 653 AA; 72053 MW; 18B07D171E874205 CRC64;

Query Match
Best Local Similarity 69.2%; Score 18; DB 1; Length 653;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXX 10
Db 173 CAAAAAAC 182

RESULT 10
ID 05H317_XANOR PRELIMINARY; PRT; 1005 AA.
AC 05H317;
DT 10-MAY-2005 (TReMBLrel. 30, Created)
DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)
DE Transcriptional regulator.
GN Name=acok; OrderedlocusNames=XO01230;
OS Xanthomonas oryzae (pv. oryzae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OC NCBI_TaxId=64187;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KACC10331 / KX085;
RX PubMed=15673718; DOI=10.1093/nar/gk1206;
RA Lee B.-M., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
RA Park I.-C., Yoon U.-H., Hahn J.-H., Koo B.-S., Lee G.-B., Kim H.,
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
RA Go S.-J.;
RT "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,
RT the bacterial blight pathogen of rice.";
RL Nucleic Acids Res. 33:577-586(2005).
DR EMBL; AE013598; AAW74484.1; -; Genomic DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:001711; P:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleic acid binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR000792; HTH_LuxR.
DR InterPro; IPR011717; TPR_4.
DR Pfam; PF00196; GcEB; 1.
DR Pfam; PF00515; TPR_1; 2.
DR Pfam; PF07721; TPR_4; 1.
DR PRINTS; PR00038; HTHLUXR.
DR ProDom; PD000307; HTH_LuxR; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00421; HTH_LuxR; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS00622; HTH_LuxR_FAMILY; 1.
DR PROSITE; PS00293; TPR_REGION; 1.
KW Complete proteome.
SQ SEQUENCE 1005 AA; 110791 MW; 662961379FD37B61 CRC64;

Query Match
Best Local Similarity 69.2%; Score 18; DB 2; Length 1005;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

Qy 1 CXXXXXXXXX 10
Db 386 CASATTAAC 395

RESULT 11
ID 040CX7_LEIMA PRELIMINARY; PRT; 1061 AA.
AC 040CX7;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Kinesin heavy chain, putative.
GN ORFNames=LmjF20_0640;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OC NCBI_TaxId=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005259; CAJ03791.1; -; Genomic DNA.
SQ SEQUENCE 1061 AA; 113930 MW; 2C69385D06371D0D CRC64;

Query Match
Best Local Similarity 69.2%; Score 18; DB 2; Length 1061;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXX 10
Db 192 CAATATSTSC 201

RESULT 12
ID 08LPD6_HORVU PRELIMINARY; PRT; 52 AA.
AC 08LPD6;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Putative glutathione-S-transferase (Fragment).
GN Name=bar2;
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OC NCBI_TaxId=4513;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Baldwin A.;
RT "The effect of thiocarbamate herbicides and the safener dichloromid in
RT barley and wild oats.";
RL Thesis (2001), Department of Cardiff School of Biosciences, Cardiff
RL university, Cardiff, United Kingdom.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Baldwin A., Harwood J.L., Machray G.C., Francis D., Rogers H.J.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ419772; CAD11963.1; -; mRNA.
DR HSSP; P12653; IAXD.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002197; HTH_Pis.
DR PRINTS; PR01590; HTHFIS.
KW Transferase.
FT NON_TER 1 1
FT NON_TER 52 52
SQ SEQUENCE 52 AA; 5750 MW; 7392EB72A3C70B9B CRC64;

```

Query Match 65.4%; Score 17; DB 2; Length 52;  
Best Local Similarity 20.0%; Pred. No. 1.3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 18 CTSASSTRSC 27

RESULT 13

Q5DCN3\_SCHJA PRELIMINARY; PRT; 122 AA.  
ID Q5DCN3\_SCHJA  
AC Q5DCN3;  
DT 10-MAY-2005 (TREMBlrel. 30, Created)  
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)  
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)  
DE Hypothetical protein.  
OS Schistosoma japonicum (Blood fluke).  
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;  
OC Schistosomatidae; Schistosomatidae; Schistosoma.  
OX NCBI\_TaxID=6182;  
RN NUCLEOTIDE SEQUENCE.  
RP Han Z.;  
RT "The full-length cDNA sequences of Schistosoma japonicum genes."  
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY814691; FAY814691.1; mRNA.  
KW Hypothetical protein.  
SQ SEQUENCE 122 AA; 13028 MW; 960E8615986958DE CRC64;

Query Match 65.4%; Score 17; DB 2; Length 122;  
Best Local Similarity 20.0%; Pred. No. 1.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 76 CTSSTSTAC 85

RESULT 14

Q96LJ4\_HUMAN PRELIMINARY; PRT; 128 AA.  
ID Q96LJ4\_HUMAN  
AC Q96LJ4;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Hypothetical protein FLJ25437.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN NUCLEOTIDE SEQUENCE.  
RP Tissue-testis;  
RC Kawakami B., Sugiyama A., Takemoto M., Suzuki Y., Hata H.,  
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,  
RA Irie K., Otsuki T., Sato H., Nishitawa T., Nagai K., Isegai T.,  
RA Sugano S.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK058166; BAB71697.1; mRNA.  
DR Ensembl; ENSG00000178093; Homo sapiens.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004672; F:protein kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR00719; Prot\_kinase.  
DR Pfam; PF00069; Kinase; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
KW ATP-binding; Kinase; Nucleotide-binding; Transferase;  
SQ SEQUENCE 128 AA; 14116 MW; D30AC8C668621B0C CRC64;

Query Match 65.4%; Score 17; DB 2; Length 128;  
Best Local Similarity 20.0%; Pred. No. 1.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 118 CTAITSATKC 127

RESULT 15

Q6ZKM7\_ORYSA PRELIMINARY; PRT; 139 AA.  
ID Q6ZKM7\_ORYSA  
AC Q6ZKM7;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Hypothetical protein OJ1118\_A06.7-1.  
GN Name=OJ1118\_A06.7-1;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriacoidae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN NUCLEOTIDE SEQUENCE.  
RP Sasaki T., Matsunoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare(GM3) genomic DNA, chromosome 8, BAC  
clone:OJ1118\_A06.7-1."  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP003873; BAD08805.1; Genomic DNA.  
DR Gramene; Q6ZKM7;  
KW Hypothetical protein.  
SQ SEQUENCE 139 AA; 14109 MW; E281360370FB63C9 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 139;  
Best Local Similarity 20.0%; Pred. No. 1.9;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
DB 77 CAITSSSSC 86

RESULT 16

Q6H482\_ORYSA PRELIMINARY; PRT; 155 AA.  
ID Q6H482\_ORYSA  
AC Q6H482;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)  
DE Hypothetical protein OJ1119\_C03.6 (Hypothetical protein  
B1012G11.50).  
GN Name=OJ1119\_C03.6; Synonyms=B1012G11.50;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriacoidae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN NUCLEOTIDE SEQUENCE.  
RP Sasaki T., Matsunoto T., Hattori M., Sasaki Y., Katayose Y.;  
RA Sasaki T., Matsunoto T., Hattori M., Sasaki Y., Katayose Y.;  
RT "Oryza sativa nipponbare(GM3) genomic DNA, chromosome 9, BAC  
clone:OJ1119\_C03.6."  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP005903; BAD38424.1; Genomic DNA.  
SQ SEQUENCE 155 AA; 14116 MW; D30AC8C668621B0C CRC64;

DR Gramene; Q6H482; -;  
KW Hypothetical protein.  
SQ SEQUENCE 155 AA; 16233 MW; 02D836E58C7B52 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 155;  
Best Local Similarity 20.0%; Pred. No. 2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXX 10  
DB 25 CQTSAADC 34

RESULT 17  
O657M0 ORYSA PRELIMINARY; PRT; 160 AA.

AC O657M0\_1  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE Hypothetical protein P0468H06.13 (Hypothetical protein P0468H05.38).  
CN Name=P0468H06.13; Synonyms=P0468H05.38;  
OS Oryza sativa (japonica cultivar-group);  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]

RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsunoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
RA Wu J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,  
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
RA Shinokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshikawa R., Yukawa K.,  
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
RA Yano M., Jiang J., Gojobori T.;  
RT "The genome sequence and structure of rice chromosome 1.";  
RL Nature 420:312-316(2002).  
DR EMBL; AP003075; BAD44997.1; -; Genomic\_DNA.  
DR EMBL; AP003267; BAD87274.1; -; Genomic\_DNA.  
DR Gramene; O657M0; -;  
KW Hypothetical protein.  
SQ SEQUENCE 160 AA; 17630 MW; C3A5ABDE5CF75DA9 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 160;  
Best Local Similarity 20.0%; Pred. No. 2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXX 10  
DB 105 CSGATSTASC 114

RESULT 18  
OAT4H1 TETNG PRELIMINARY; PRT; 160 AA.

AC OAT4H1\_1  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Chromosome undetermined SCAP959, whole genome shotgun sequence.  
CN ORFname=GSTENG00007326001;  
OS Tetraodon nigroviridis (Green puffer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;  
OC Tetraodontidae; Tetraodontidae; Tetraodon.  
OX NCBI\_TaxID=99883;  
RN [1]

RP NUCLEOTIDE SEQUENCE.  
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,  
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
RA Anthouard V., Jubin C., Caetelli V., Karinka M., Vacherie B.,  
RA Biemont C., Skalli Z., Catcolico L., Poulain J., De Bernardis V.,  
RA Cruaud C., Duprat S., Broctier P., Coutanceau J.P., Gouzy J.,  
RA Parra G., Lardier S., Chapelle C., McKernan K.J., McEwan P., Bosak S.,  
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
RA Laudet V., Schachter V., Queller F., Saurin W., Scarpelli C.,  
RA Wincker P., Lander E.S., Weissbach J., Roest Crolians H.;  
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
RT the early vertebrate proto-karyotype.";  
RL Nature 431:946-957(2004).  
RN [2]

RP NUCLEOTIDE SEQUENCE.  
RG Genoscope, Whitehead Institute Centre for Genome Research;  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; CAAB0100659; CAP92211.1; -; Genomic DNA.  
SQ SEQUENCE 160 AA; 17094 MW; 6167384F2A5E51B CRC64;

Query Match 65.4%; Score 17; DB 2; Length 160;  
Best Local Similarity 20.0%; Pred. No. 2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXX 10  
DB 115 CTTAASSTSC 124

RESULT 19  
O825B8 STRAW PRELIMINARY; PRT; 219 AA.  
ID O825B8;  
AC O825B8;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Putative tetr-family transcriptional regulator.  
GN Ordered locus names=SAV7510;  
OS Streptomyces avermitilis;  
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;  
OC Streptomycetaceae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=33903;  
RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RA STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RC MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;  
RA Shinoe S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
RA Shinose M., Takahashi Y., Horikawa H., Nakawara H., Osone T.,  
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,  
RT "Genome sequence of an industrial microorganism Streptomyces  
RT avermitilis: deducing the ability of producing secondary  
RT metabolites.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RC MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;  
RA Ikeda H., Ishikawa J., Hanamoto A., Shinoe M., Kikuchi H., Shiba T.,  
RA Sakaki Y., Hattori M., Omura S.;  
RT "Complete genome sequence and comparative analysis of the industrial  
RT microorganism Streptomyces avermitilis.";  
RT Nat. Biotechnol. 21:526-531(2003).  
DR EMBL; BA000030; BAC75221.1; -; Genomic\_DNA.

```
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR01647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PROSITE; PS50977; HTH_TETR_2; 1.
KW Complete proteome.
SQ SEQUENCE 219 AA; 24066 MW; 5C59ADP39430F304 CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 219;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 151 CAEAAAAAAC 160

RESULT 20
O6152_9VIRU PRELIMINARY; PRT; 220 AA.
ID O6152_9VIRU PRELIMINARY; PRT; 220 AA.
AC O66152;
DT 01-JUN-2003 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Coat protein.
OS Cymbidium mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flexiviridae;
OC Potexvirus.
OX NCBI_TaxId=12178;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Korean isolate;
RX MEDLINE=95278762; PubMed=7758973; DOI=10.1016/0378-1119(95)00105-F;
RA Ryu K.H., Yoon K.B., Park W.M.;
RT "Nucleotide sequence of coat protein gene of cymbidium mosaic
RT potexvirus genomic RNA, the Korean isolate.";
RL Gene 156:303-304(1995).
DR EMBL; X81051; CAAS6941.1; -; Genomic_RNA.
DR PIR; JC4082; JC4082.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000052; P1vir coat.
DR Pfam; PF00286; ViruB_P-coat_1.
DR PRINTS; PR00232; POTCARLCOAT.
DR ProDom; PD00603; P1vir_coat; 1.
KW Capsid protein.
SQ SEQUENCE 220 AA; 23719 MW; 1F40C6E7CDCCF3B8 CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 220;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 166 CSAATLTATC 175

RESULT 21
O6YTS2_ORYSA PRELIMINARY; PRT; 256 AA.
ID O6YTS2_ORYSA PRELIMINARY; PRT; 256 AA.
AC O6YTS2_
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein P0419H09.29.
GN Name=P0419H09.29;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxId=39947;
RN [1]
```

```
RP NUCLEOTIDE SEQUENCE.
RA Saeki T., Matsunoto T., Katayose Y.;
RL Submitted (NOV-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP005918; BAD10719.1; -; Genomic_DNA.
DR Gramene; O6YTS2; -.
KW Hypothetical protein.
SQ SEQUENCE 256 AA; 26591 MW; 63BFCAF53D2D40EA CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 256;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
DB 203 CSASASARC 212

RESULT 22
O65023_ORYSA PRELIMINARY; PRT; 260 AA.
ID O65023_ORYSA PRELIMINARY; PRT; 260 AA.
AC O65023;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE DnaJ-like protein (Putative heat shock protein).
GN Name=OSUNB0027B08.12; Synonyms=OSUNB0078D06.16;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxId=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tselir T., Kim M.M., Bera J.V., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldlyum T.V.,
RA Yang Q.O., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (MAY-2003) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: Have a continuous role in plant development probably in
CC the structural organization of compartments (By similarity).
DR EMBL; AC133778; MA039854.1; -; Genomic_DNA.
DR EMBL; AC133339; AAP4230.1; -; Genomic_DNA.
DR HSSP; P25685; 1HDJ.
DR Gramene; O65023; -.
DR GO; GO:0031072; F:heat shock protein binding; IEA.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR GO; GO:0006986; P:response to unfolded protein; IEA.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ_1.
DR PRINTS; PR00625; DNAJPROTEIN.
DR SMART; SM00271; DnaJ_1.
DR PROSITE; PS00636; DNAJ_1; 1.
DR PROSITE; PS50076; DNAJ_2; 1.
KW Chaperone; Heat shock.
SQ SEQUENCE 260 AA; 28639 MW; DDC645DAF9BA1ED CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 260;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXC 10
```

Db 19 CAASAFSAC 28

## RESULT 23

Q57VE4\_ANOGA PRELIMINARY; PRT; 304 AA.  
AC Q57VE4;  
DT 01-FEB-2005 (TREMBlrel. 29, Created)  
DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)  
DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)  
DE ENSANGP0000025853 (Fragment).  
GN ORFNames=ENSANG0000024781;  
OS Anopheles gambiae str. PEST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;  
OC Anophelinae; Anopheles.  
OX NCBI\_Taxid=180454;  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=PEST;  
RC The Anopheles gambiae Sequence Committee;  
RT "Anopheles gambiae re-annotation."  
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=PEST;  
RG The Anopheles gambiae Sequence Committee;  
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.  
CC -! CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AAB01008839; EAL41365.1; -; Genomic\_DNA.  
FT NON\_TER 1 304  
SQ SEQUENCE 304 AA; 29530 MW; 1F55F0FB846B16B9 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 304;  
Best Local Similarity 20.0%; Pred. No. 2.6;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10  
Db 20 CSSTSTTVC 29

## RESULT 24

Q582B8\_9TRYP PRELIMINARY; PRT; 344 AA.  
AC Q582B8;  
DT 10-MAY-2005 (TREMBlrel. 30, Created)  
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)  
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=TP927.5.2480;  
OS Trypanosoma brucei.  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_Taxid=5631;  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=GUTat10.1;  
RC Ghedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,  
RA Shalom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,  
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,  
RA Khalak H.G., Salzberg S., Simpson A.U., Tallon L., Van Aken S.,  
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;  
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=GUTat10.1;  
RA El-Sayed N.M., Khalak H., Adams M.D.;  
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=GUTat10.1;  
RA Haas B., Blandin G., El-Sayed N.;  
RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC091655; AAX80451.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 344 AA; 37545 MW; A3925C9B2A48BD53 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 344;  
Best Local Similarity 20.0%; Pred. No. 2.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10  
Db 251 CSSTASGSC 260

## RESULT 25

Q57Z20\_9TRYP PRELIMINARY; PRT; 361 AA.  
AC Q57Z20;  
DT 10-MAY-2005 (TREMBlrel. 30, Created)  
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)  
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=TP927.5.4290;  
OS Trypanosoma brucei.  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_Taxid=5631;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=GUTat10.1;  
RA Ghedin E., Blandin G., Bartholomew D., Caler E., Haas B., Hannick L.,  
RA Shalom J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,  
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,  
RA Khalak H.G., Salzberg S., Simpson A.U., Tallon L., Van Aken S.,  
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;  
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=GUTat10.1;  
RA El-Sayed N.M., Khalak H., Adams M.D.;  
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=GUTat10.1;  
RA Haas B., Blandin G., El-Sayed N.;  
RL Submitted (Apr-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC116668; AAX80579.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 361 AA; 40083 MW; 95BC9BDCB3B9895 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 361;  
Best Local Similarity 20.0%; Pred. No. 2.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXXC 10  
Db 205 CSSTASSSC 214

## RESULT 26

Q4YW87\_PLABE PRELIMINARY; PRT; 364 AA.  
AC Q4YW87;  
DT 13-SEP-2005 (TREMBlrel. 31, Created)  
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)  
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)  
DE Hypothetical protein (Fragment).  
GN ORFNames=PB105184.00.0;  
OS Plasmodium berghei.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_Taxid=5821;

[1]  
RN NUCLEOTIDE SEQUENCE.  
RP Hall N., Kairas M., Raine J.D., Carlton J.M., Kooij T.W.A.,  
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,  
RA James K., Rutherford K., Harris B., Trueman H.E., Mendoza J.,  
RA Quill M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,  
RA Bidwell S.L., Rajandream M.A., Carnocci D.J., Yates J.R., Kafatos F.C.,  
RA Jance C.J., Barrett B., Turner C.M.R., Waters A.P., Sinden R.S.,  
RT "A comprehensive survey of the Plasmodium life cycle by genomic,  
transcriptomic, and proteomic analyses."  
RL Science 307:82-86(2005).  
CC -!- CAUTION: The sequence shown here is derived from an  
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
preliminary data.  
CC EMBL: CA101002130; CAH97719.1; -; Genomic\_DNA.  
KM Hypothetical protein.  
FT NON TER 1  
SQ SEQUENCE 364 AA; 41304 MW; 1C1711BE940AA94F CRC64;  
QY Query Match 65.4%; Score 17; DB 2; Length 364;  
Best Local Similarity 20.0%; Pred. No. 2.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
1 CXXXXXXXXX 10  
DB 332 CSTSTTASCS 341  
QY 1 CXXXXXXXXX 10  
DB 332 CSTSTTASCS 341  
RESULT 27  
O7ORJ3 GERHY PRELIMINARY; PRT; 368 AA.  
ID O7ORD3;  
AC O7ORD3;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE MYB1 protein.  
CN Name=Myb1;  
OS Gerbera hybrid cv. 'Terra Regina'.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
OC asterids; campanulids; Asterales; Asteraceae; Mutisioideae; Mutisieae;  
OC Gerbera.  
OC NCBI\_TaxID=226891;  
OX [1]  
RN NUCLEOTIDE SEQUENCE.  
RP TISSUE=Petal;  
RC PubMed=14605235; DOI=10.1104/pp.103.026039;  
RX Elomaa P., Uinari A., Mehto M., Albert V.A., Laitinen R., Teeri T.H.,  
RA "Activation of anthocyanin biosynthesis in Gerbera hybrida  
(Asteraceae) suggests conserved protein-protein and protein-promoter  
interactions between the anciently diverged monocots and eudicots."  
RT Plant Physiol. 133:1831-1842(2003).  
RL -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC EMBL: AJ554697; CAD87007.1; -; mRNA.  
DR HSSP; Q03237; IASJ.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0005677; F:DNA binding; IEA.  
DR InterPro; IPR012287; Homeodomain-rel.  
DR InterPro; IPR010005; Myb DNA bd.  
DR Pfam; PF00249; Myb DNA-binding; 2.  
DR SMART; SM00717; SANT; 2.  
DR PROSITE; PS00037; MYB\_1; UNKOWN\_1.  
DR PROSITE; PS00334; MYB\_2; 1.  
DR PROSITE; PS50090; MYB\_3; 2.  
DR PROSITE; PS50090; MYB\_3; 2.  
KM Nuclear protein; Repeat.  
SQ SEQUENCE 368 AA; 40911 MW; 7CFEFCBAD3AD9421 CRC64;  
QY Query Match 65.4%; Score 17; DB 2; Length 368;  
Best Local Similarity 20.0%; Pred. No. 2.7;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
1 CXXXXXXXXX 10

DB 286 CATSTSTDC 295  
RESULT 28  
ID CBID\_AGRTS STANDARD; PRT; 376 AA.  
AC 08UB06;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Putative cobalt-precursorin-6A synthase [deacylating] (EC 2.1.1.-).  
GN Name=cbid; OrderedLocustNames=Atu2795, AGP C 5073;  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Agrobacterium.  
OX NCBI\_TaxID=176299;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;  
RA Wood D.W., Secubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,  
RA Chen Y., Paulsen I.T., Bisen J.A., Karp P.D., Boyee D., St.,  
RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,  
RA Kutyavina T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,  
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krepan W., Perry M.,  
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
RA Nester E.W.;  
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
C58."  
RT Science 294:2317-2323(2001).  
RL [2]  
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RP MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;  
RX Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
RA Gurello B., Goldman B.S., Cao Y., Askenazi M., Halling L.,  
RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
RA Wollam C., Allinger M., Doughy D., Scott C., Jappes C., Markelz B.,  
RA Flanagan C., Crowell C., Garson J., Lomo C., Sear C., Strub G.,  
RA Cielo C., Slater S.;  
RT "Genome sequence of the plant pathogen and biotechnology agent  
Agrobacterium tumefaciens C58."  
RT Science 294:2323-2328(2001).  
RL -!- FUNCTION: May catalyze the methylation of C-1 in cobalt-precursorin-  
5 and the subsequent extrusion of acetic acid from the resulting  
intermediate to form cobalt-precursorin-6A.  
CC -!- PATHWAY: Adenosylcobalamin biosynthesis; anaerobic branch of  
corrin ring synthesis; seventh step.  
CC -!- SIMILARITY: Belongs to the cbid family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC EMBL; AE009227; AAL45776.1; -; Genomic\_DNA.  
DR EMBL; AE008192; AAK8508.1; ALT\_INIT; Genomic\_DNA.  
DR PIR; AB2920; AB2920.  
DR PIR; C97694; C97694.  
DR HAMAP; MF\_00787; -; 1.  
DR InterPro; IPR002748; Cbid.  
DR Pfam; PF01888; Cbid; 1.  
DR TIGRFAMs; TIGR00312; cbid; 1.  
KM Cobalamin biosynthesis; Complete proteome; Transferase.  
SQ SEQUENCE 376 AA; 39088 MW; D5437CB4CA463D67 CRC64;  
QY Query Match 65.4%; Score 17; DB 1; Length 376;  
Best Local Similarity 20.0%; Pred. No. 2.8;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
1 CXXXXXXXXX 10



RA Cheepel Y., Collamore A., Cook A., Cooke P., Corum B., Dearlano K.,  
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,  
 RA Erickson J., Fato S., Ferreira P., FitzGerald M., Gage D., Galagan J.,  
 RA Gardyna S., Gierre S., Graham L., Grand-Pierre N., Hatz N.,  
 RA Hagoopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,  
 RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,  
 RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,  
 RA Ma L.-Y., Mabbitt R., Maclean C., MacDonald P., Major J., Manning J.,  
 RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,  
 RA Milnova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,  
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill S.,  
 RA Oliver J., Peterson K., Phunkham P., Pierre N., Purcell S.,  
 RA Rachupa A., Ramsamy U., Raymond C., Reta R., Rise C., Rogov P.,  
 RA Roman J., Schauer S., Schupack R., Seaman S., Severy P., Smirnov S.,  
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,  
 RA Talamas J., Testaye S., Theodore J., Topham K., Travers M.,  
 RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,  
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,  
 RA Lander E.,  
 RT "Fusarium graminearum genome sequence."  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL, AACM01000320; EAA77994.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 437 AA; 47817 MW; 01EBA0260503394 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 437;  
 Query Local Similarity 20.0%; Pred. No. 2.9;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CXXXXXXC 10  
 Db 199 CATLASTSSC 208

RESULT 33  
 Q9VMG7 DROME PRELIMINARY; PRT; 471 AA.  
 ID Q9VMG7 DROME PRELIMINARY; PRT; 471 AA.  
 AC Q9VMG7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE CG13990-PA.  
 GN Name=CG13990; ORFNames=CG13990;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydriidae; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Doudon K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostel A., Gabrielian A.B., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke J., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Liao P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Metel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulyov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon S., Nuskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195 (2000).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celinker S.E., Wheeler D.A., Kromliller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence."  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kromliller B., Carlson J.W., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celinker S.B.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomic perspective."  
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20 (2002).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Miya S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tudy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Betencourt B.R., Celinker S.E., de Grey A.D.N.-J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review."  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Berkeley Drosophila Genome Project;  
 RA Celinker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,  
 RA Hoskins R., Stapleton M., Pacle J., Park S., Svirskas R., Smith E.,  
 RA Yu C., Rubin G.;  
 RT "Drosophila melanogaster release 4 sequence."  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Flybase;  
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AEO03612; AAF52351.2; -; Genomic DNA.  
 DR Ensemble; CG13990; Drosophila melanogaster.  
 DR Flybase; FBgn040950; CG13990.  
 DR GO; GO:0005576; C:extracellular region; IEA.  
 DR GO; GO:0008061; P:chitin binding; IEA.  
 DR GO; GO:0006030; P:chitin metabolism; IEA.  
 DR InterPro; IPR002557; Chitin\_bind\_Pera.  
 DR Pfam; PF01607; CBM\_14; 2.



DR SMART; SM00494; ChtBD2; 2.  
DR PROSITE; PSS0940; CHIT\_BIND II; 2.  
SQ SEQUENCE 471 AA; 4811 MW; F3167E80728E9578 CRC64;  
Query Match 65.4%; Score 17; DB 2; Length 471;  
Best Local Similarity 20.0%; Pred. No. 3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
OY 1 CXXXXXXXXX 10  
DB 350 CASSSTTTC 359  
RESULT 34  
O7X744 ORYSA PRELIMINARY; PRT; 475 AA.  
ID O7X744\_ORYSA PRELIMINARY; PRT; 475 AA.  
AC O7X744;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)  
DE OSJNB0060P14.14 protein (OSJNB0048E02.5 protein).  
CN Name=OSJNB0060P14.14; Synonyms=OSJNB0048E02.5;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=39947;  
RN NUCLEOTIDE SEQUENCE.  
RP PubMed=12447439; DOI=10.1038/nature01183;  
RX Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,  
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,  
RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Pan D.,  
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Guan J.,  
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,  
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Yu Y., Jia J., Zhang Y.,  
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,  
RA Lan L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,  
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,  
RA Han B.;  
RT "Sequence and analysis of rice chromosome 4.";  
RL Nature 420:316-320(2002).  
DR EMBL; AL663017; CAE04361.2; -; Genomic DNA.  
DR EMBL; AL606653; CAE04825.2; -; Genomic DNA.  
DR Gramene; O7X744; -;  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR001810; F-box.  
DR Pfam; PF00646; F-box; 1.  
DR SMART; SM00256; FBOX; 1.  
DR PROSITE; PSS0181; FBOX; 1.  
SQ SEQUENCE 475 AA; 53063 MW; 63345B3CB2D4CA5 CRC64;  
Query Match 65.4%; Score 17; DB 2; Length 475;  
Best Local Similarity 20.0%; Pred. No. 3;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
OY 1 CXXXXXXXXX 10  
DB 417 CSSSTLSSAC 426  
RESULT 35  
O4P4C7 USTMA PRELIMINARY; PRT; 491 AA.  
ID O4P4C7\_USTMA PRELIMINARY; PRT; 491 AA.  
AC O4P4C7;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=UM05036.1;  
OS Ustilago maydis 521.  
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;  
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.

OX NCBI\_TaxID=237631;  
RN NUCLEOTIDE SEQUENCE.  
RP PubMed=521;  
RX Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,  
RA Alt-Zandra M., Allen N., Allen T., An P., Anderson M., Anderson S.,  
RA Arachchi H., Armbruster J., Bachantang P., Baldwin J., Barry A.,  
RA Bayat T., Blitshteyn B., Bloom T., Blye J., Boguslavsky L.,  
RA Borowsky M., Boukhalter B., Brumache A., Butler J., Caliste N.,  
RA Calvo S., Camarata J., Campo K., Chang J., Cheshtang Y., Citroen M.,  
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,  
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,  
RA Dorjee K., Dorris L., Dufley N., Dupes A., Ekins T., Engels R.,  
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,  
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,  
RA Gritke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,  
RA Hagopian D., Hages B., Hall J., Hatcher B., Heller A., Higgins H.,  
RA Honan T., Horn A., Houde N., Hughes L., Huine W., Husby E., Iliev I.,  
RA Jaffe D., Jones C., Kamel M., Kamet A., Kamyssele M., Karlsson E.,  
RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutti K.,  
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,  
RA Lindblad-Toh K., Liu X., Lokitsang T., Lokitsang Y., Lucien O.,  
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,  
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,  
RA McCarthy M., McDonough S., McInnes T., Meldrum J., Meneus L.,  
RA Mesirov J., Mhallev A., Minova T., Mikkelsen T., Mieng V., Moru K.,  
RA Mozes J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,  
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,  
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omorcho B.,  
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piganì B.,  
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,  
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,  
RA Rutman M., Schubach R., Seaman C., Settipalli S., Sharpe T.,  
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnuez C.,  
RA Spencer B., Stalker J., Strange-Thomann N., Stavropoulos S.,  
RA Stenson P., Stone C., Stone S., Stubbs M., Talamas Y., Tchuinga P.,  
RA Teasing P., Testaye S., Theodore J., Thoulustang Y., Topham K.,  
RA Towey S., Tsamila T., Tsomo N., Vallée D., Vassiliev H.,  
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,  
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,  
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zairoun J., Zembek L.,  
RA Zimmer A., Zody M., Zander B.;  
RT "The genome sequence of Ustilago maydis.";  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AACP0100186; EAK85896.1; -; Genomic DNA.  
DR Hypothetical protein.  
KW KX  
SQ SEQUENCE 491 AA; 51869 MW; 768A84FB8BD71BBE CRC64;  
Query Match 65.4%; Score 17; DB 2; Length 491;  
Best Local Similarity 20.0%; Pred. No. 3.1;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
OY 1 CXXXXXXXXX 10  
DB 453 CAASTGSSSC 462  
RESULT 36  
O4XZV5 PLACH PRELIMINARY; PRT; 520 AA.  
ID O4XZV5\_PLACH PRELIMINARY; PRT; 520 AA.  
AC O4XZV5;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein (Fragment).  
GN ORFNames=PC000338.02.0;  
OS Plasmodium chabaudi.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5825;  
RN NUCLEOTIDE SEQUENCE.

RP NUCLEOTIDE SEQUENCE.  
RA Hall N., Karris M., Raine J.D., Carlton J.M., Kooij T.M.A.,  
RA Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,  
RA James K., Rutledge K., Harris B., Harris D., Church C.,  
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,  
RA Bidwell S.L., Rajandream M.A., Carnucci D.J., Yates J.R., Kafatos F.C.,  
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.,  
RT "A comprehensive survey of the plasmidium life cycle by genomic,  
RT transcriptomic, and proteomic analyses."  
RL Science 307:82-86(2005).  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
CC EMBL: CAJ01002036; CAH77555.1; -: Genomic\_DNA.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00400; WD40; 5.  
DR PRINTS: PR00320; GPROTEINBRPT.  
DR SMART: SM00320; WD40; 5.  
DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_1.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Hypothetical protein; Repeat; WD repeat.  
FT NON\_TER 520 520  
SQ SEQUENCE 520 AA; 59951 MW; FA789C11BFD99087 CRC64;  
  
Query Match 65.4%; Score 17; DB 2; Length 520;  
Best Local Similarity 20.0%; Pred. No. 3.1;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXC 10  
DB 364 CSTSLSSSC 373  
  
RESULT 37  
Q0803D5 BRARE PRELIMINARY; PRT; 556 AA.  
AC Q0803D5  
ID Q0803D5  
DT 01-JUN-2003 (TRENBLREL. 24, Created)  
DT 01-JUN-2003 (TRENBLREL. 24, Last sequence update)  
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)  
DE Similar to adrenal secretory serine protease.,  
GN ORFNames=zgc:55888;  
OS Brachydanio rerio (zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN-AB; TISSUE=whole body;  
RX MEDLINE=22388557; PubMed=12477932; DOI=10.1073/pnae.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,  
RA Alechul S.F., Zeeberg B., Buettow K.H., Scheffer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stabileton M., Soares M.B., Bonaldo M.F., Caesavart T.L., Schetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rahn S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettlemen M., Madan A., Rodighiero S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
RA Scheraga A., Schein J.E., Jones S.J.M., Maira M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.

RC STRAIN-AB; TISSUE=whole body;  
RA Strausberg R.,  
RT Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC044526; AAH44526.1; -: mRNA.  
DR HSSP: P00760; IEZK.  
DR Ensembl: ENSDARG0000016538; Danio rerio.  
DR ZFIN: ZDB-GENE-040426-834; zgc:55888.  
DR GO: GO:0004263; F:chymotrypsin activity; IEA.  
DR GO: GO:0008233; F:peptidase activity; IEA.  
DR GO: GO:0044295; F:trypsin activity; IEA.  
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro: IPR001254; Peptidase\_S1\_S6.  
DR InterPro: IPR001314; Peptidase\_S1A.  
DR Pfam: PF00089; Trypsin\_2.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR SMART: SM00020; TRYP\_SPC; 2.  
DR PROSITE: PS50240; TRYP\_SIN\_DOM; 2.  
DR PROSITE: PS00134; TRYP\_SIN\_HIS; UNKNOWN\_1.  
DR PROSITE: PS00135; TRYP\_SIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
SQ SEQUENCE 556 AA; 61077 MW; C7B23D930547878C CRC64;  
  
Query Match 65.4%; Score 17; DB 2; Length 556;  
Best Local Similarity 20.0%; Pred. No. 3.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXC 10  
DB 493 CTHAAASTSC 502  
  
RESULT 38  
Q05GXM1\_XANOR PRELIMINARY; PRT; 568 AA.  
AC Q05GXM1  
ID Q05GXM1  
DT 10-MAY-2005 (TRENBLREL. 30, Created)  
DT 10-MAY-2005 (TRENBLREL. 30, Last sequence update)  
DT 10-MAY-2005 (TRENBLREL. 30, Last annotation update)  
DE Alkaline phosphatase.  
GN Name=phoA; OrderedLocustNames=X003296;  
OS Xanthomonas oryzae (pv. oryzae).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xanthomonas.  
OX NCBI\_TaxID=64187;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=KACC10331 / KX085;  
RX PubMed=15673718; DOI=10.1093/nar/gk1206;  
RA Lee B.-M., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,  
RA Park I.-C., Yoon U.-H., Hahn J.-H., Ko B.-S., Lee G.-B., Kim H.,  
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,  
GO S.-J.;  
RT "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,  
RT the bacterial blight pathogen of rice.";  
RL Nucleic Acids Res. 33:577-586(2005).  
RU EMBL: AE013598; AAW7550.1; -: Genomic\_DNA.  
DR GO: GO:0008152; P:metabolism; IEA.  
DR InterPro: IPR001952; ALK\_phosphatase.  
DR Pfam: PF00245; ALK\_phosphatase.  
DR PRINTS: PR00113; ALKPHPTASE.  
DR SMART: SM00098; alkppc; 1.  
KW Complete proteome.  
SQ SEQUENCE 568 AA; 60106 MW; CAE69E266ACB3A6 CRC64;  
  
Query Match 65.4%; Score 17; DB 2; Length 568;  
Best Local Similarity 20.0%; Pred. No. 3.2;  
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXC 10  
DB 21 CASTAGSSAC 30

```

RESULT 39
QAN2Y7_THEMA PRELIMINARY; PRT; 626 AA.
ID QAN2Y7_THEMA PRELIMINARY;
AC QAN2Y7;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein.
OS ORFNames=TF04_0205;
ON Theileria parva.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileriidae;
OC Theileria.
OX NCBI_TaxID=5875;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Muguga.
RA Gardner M.J., Bishop R., Shah T., de Villiers E.P., Carlton J.M.,
RA Hall N., Ren Q., Paulsen I.T., Pain A., Berriman M., Wilson R.J.M.,
RA Saco S., Ralph S.A., Mann D.J., Xiong Z., Shallom S.J., Weidman J.,
RA Jiang L., Lynn J., Weaver B., Shoibh A., Wasawo D., Crabtree J.,
RA Wortman J.C., Utterback T.R., Feldblyum T.V., Petrea M., Allen J.,
RA Silva J.C., Utterback T.R., Feldblyum T.V., Petrea M., Allen J.,
RA Taracha E.L.N., Salzberg S.L., White O.R., Fitzhugh H.A., Morzaria S.,
RA Venter J.C., Fraser C.M., Nene V.;
RT "Genome sequence of Theileria parva, a bovine pathogen that transforms
RT lymphocytes."
RT Science 309:134-137(2005) .
RL [2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Muguga.
RA Gardner M., Bishop R., Shah T., de Villiers E., Carlton J.M., Hall N.,
RA Ren Q., Paulsen I.T., Pain A., Berriman M., Wilson R.J.M., Sato S.,
RA Ralph S.A., Mann D.J., Xiong Z., Shallom S.J., Weidman J., Jiang L.,
RA Lynn J., Weaver B., Shoibh A., Wasawo D., Crabtree J., Wortman J.R.,
RA Haas B., Anguilo S., Creasy T.H., Lu C., Suh B., Silva J.C.,
RA Utterback T., Feldblyum T., Petrea M., Allen J., Taracha E.L.,
RA Salzberg S.L., White O., Fitzhugh H.A., Morzaria S., Venter J.C.,
RA Fraser C.M., Nene V.;
RA Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAGK01000004; EANJ1557.1; -; Genomic DNA.
CC DR Hypothetical protein.
CC KW SEQUENCE 626 AA; 72615 MW; 663D85BA95D09916 CRC64;
SQ
Query Match: 65.4%; Score 17; DB 2; Length 626;
Best Local Similarity 20.0%; Pred. NO. 3.4;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 1 CXXXXXXXXXC 10
DB 44 CAASSDASC 53

```

RT	clone:OSUNBA0029G06."				
RL	submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.				
RN	[2]				
RP	NUCLEOTIDE SEQUENCE.				
RA	Sasaki T., Matsumoto T., Yamamoto K.;				
RT	"Oryza sativa nphosphabare(GA3) genomic DNA, chromosome 6, BAC				
RT	clone:OSUNBA0072A21."				
RL	Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL; AF004680; BAD35689.1; -; Genomic DNA.				
DR	Graeme; Q697Y8; -;				
DR	GO; GO:0005524; F:ATP binding; IEA.				
DR	GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.				
DR	GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.				
DR	GO; GO:0016998; P:cell wall catabolism; IEA.				
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.				
DR	InterPro; IPR002482; LysM.				
DR	InterPro; IPR000719; Prot_kinase.				
DR	InterPro; IPR002290; Ser_Thr_kinase.				
DR	InterPro; IPR01245; Tyr_kinase.				
DR	Pfam; PF01476; LysM; 1.				
DR	Pfam; PF00069; Kinase; 1.				
DR	ProDom; PD000001; Prot_kinase; 1.				
DR	SMART; SM00220; S_TKC; 1.				
DR	SMART; SM00219; TYKc; 1.				
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.				
KW	Kinase; Receptor.				
SO	SEQUENCE 630 AA; 67017 MW; FOEFGBB8A63646 CRC64;				
Query Match	65.4%; Score 17; DB 2; Length 630;				
Bee Local Similarity	20.0%; Pred. No. 3.4;				
Matches	2; Conservative	0; Mismatches	8; Indels	0; Gaps	0;
Oy	1 CXXXXXXXXC 10				
Dd	48 CSERTATTSC 57				
RESULT 41					
HSP71 LEIMA					
ID	HSP71 LEIMA STANDARD; PRT; 634 AA.				
AC	P12076;				
DT	01-OCT-1989 (Rel. 12, Created)				
DT	01-FEB-2005 (Rel. 46, last sequence update)				
DT	13-FEB-2005 (Rel. 48, last annotation update)				
DE	Heat shock 70-related protein 1, mitochondrial precursor.				
CN	Name=HSP70.1;				
CS	Leishmania major.				
OC	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.				
OX	NCBI_TaxID=5664;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=MOM/IL/81/Friedlin;				
RA	MEDLINE=93300981; PubMed=8314893;				
RX	Searle S., McCrossan M.V., Smith D.F.;				
RT	"Expression of a mitochondrial stress protein in the protozoan				
RT	parasite leishmania major.";				
RL	J. Cell Sci. 104:1091-1100(1993).				
RN	[2]				
RP	SEQUENCE REVISION TO 461-500.				
RA	Smith D.F.;				
RL	submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.				
RN	[3]				
RP	NUCLEOTIDE SEQUENCE OF 1-249.				
RC	STRAIN=MOM/IL/81/Friedlin;				
RA	MEDLINE=89345072; PubMed=2762121;				
RX	Searle S., Campos A.J.R., Coulson R.M.R., Spithill T.W., Smith D.F.;				
RT	"A family of heat shock protein 70-related genes are expressed in the				
RT	promastigotes of leishmania major.";				
RL	Nucleic Acids Res. 17:5081-5095(1989).				
CC	-1- SUBCELLULAR LOCATION: Mitochondrial.				
CC	-1- SIMILARITY: Belongs to the heat shock protein 70 family.				

```
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X64137; CAA5498.2; -; Genomic_DNA.
DR EMBL; X14574; CAA32713.1; -; Genomic_DNA.
DR PIR; S33575; S33575.
DR HSSP; P04475; 1DKG.
DR InterPro; IPR012725; Dnak_prot.
DR InterPro; IPR001023; Hsp70.
DR PANTHER; PTHR19375; Hsp70; 1.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR TIGRFAMs; TIGR02350; prok_dnak; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; FALSE_NEG.
DR ATP-binding; Coiled coil; Heat shock; Mitochondrion; Multigene family;
DR Nucleotide-binding; Transist peptide.
FT TRANSIT 1 634 Mitochondrion (By similarity).
FT CHAIN 21 634 Heat shock 70-related protein 1.
FT COLLED 538 614 Potential.
SQ SEQUENCE 634 AA; 68330 MW; 8382B1C40C108EEA CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 1; Length 634;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
Db 7 CGSAAASAAAC 16

RESULT 42
Q0747_LEIMA
ID Q0747_LEIMA PRELIMINARY; PRT; 635 AA.
AC Q0747;
DT 13-SEP-2005 (TREMblrel. 31, Created)
DT 13-SEP-2005 (TREMblrel. 31, Last sequence update)
DE 13-SEP-2005 (TREMblrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=LmjF30.2460;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Frledlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL; CT005267; CAJ06531.1; -; Genomic_DNA.
DR InterPro; IPR012725; Dnak_prot.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR ProDom; PD000089; Hsp70; 1.
DR TIGRFAMs; TIGR02350; prok_dnak; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
DR ATP-binding; Heat shock; Nucleotide-binding.
SQ SEQUENCE 635 AA; 68948 MW; 1A11F63117CA2200 CRC64;

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 635;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 CXXXXXXXXXC 10
Db 7 CGSAAASAAAC 16

RESULT 43
Q0740_LEIMA
ID Q0740_LEIMA PRELIMINARY; PRT; 652 AA.
AC Q0740;
DT 13-SEP-2005 (TREMblrel. 31, Created)
DT 13-SEP-2005 (TREMblrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMblrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=LmjF30.2550;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Frledlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL; CT005267; CAJ06539.1; -; Genomic_DNA.
DR InterPro; IPR012725; Dnak_prot.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.

Query Match
Best Local Similarity 65.4%; Score 17; DB 2; Length 652;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
Db 7 CGSAAASAAAC 16

RESULT 44
Q0744_LEIMA
ID Q0744_LEIMA PRELIMINARY; PRT; 660 AA.
AC Q0744;
DT 13-SEP-2005 (TREMblrel. 31, Created)
DT 13-SEP-2005 (TREMblrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMblrel. 31, Last annotation update)
DE Heat shock 70-related protein 1, mitochondrial, putative.
GN ORFNames=LmjF30.2490;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Frledlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.
DR EMBL; CT005267; CAJ06539.1; -; Genomic_DNA.
DR InterPro; IPR012725; Dnak_prot.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; HSP70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
```

DR ProDom; PD000089; Hsp70; 1.  
 DR TIGRPFAMs; TIGR02350; prok\_dnaK; 1.  
 DR PROSITE; PS00297; HSP70\_1; 1.  
 DR PROSITE; PS00329; HSP70\_2; 1.  
 DR PROSITE; PS01036; HSP70\_3; 1.  
 KW ATP-binding; Heat shock; Nucleotide-binding.  
 SQ SEQUENCE 660 AA; 71621 MW; 5E1D306B3D524242 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 660;  
 Best Local Similarity 20.0%; Pred. No. 3.4;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
 Db 7 CGSAAASAC 16

RESULT 45  
 ID Q40745 LEIMA PRELIMINARY; PRT; 662 AA.

AC Q40745\_1  
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)  
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)  
 DE Heat shock 70-related protein 1, mitochondrial, putative.  
 GN ORFNames=LmjP30.2470, LmjP30.2480;  
 OS Leishmania major.

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_TaxID=5664;  
 RN (1)

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Friedlin;

RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,  
 RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,  
 RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;  
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the heat shock protein 70 family.

DR EMBL; C0005267; CAJ06535.1; -; Genomic DNA.  
 DR EMBL; C0005267; CAJ06535.1; -; Genomic DNA.

DR InterPro; IPR012725; DnaK\_prok.

DR InterPro; IPR001023; Hsp70.

DR Pfam; PF00012; HSP70; 1.

DR ProDom; PD000089; Hsp70; 1.

DR ProDom; PD000089; Hsp70; 1.

DR TIGRPFAMs; TIGR02350; prok\_dnaK; 1.

DR PROSITE; PS00297; HSP70\_1; 1.

DR PROSITE; PS00329; HSP70\_2; 1.

DR PROSITE; PS01036; HSP70\_3; 1.

KW ATP-binding; Heat shock; Nucleotide-binding.

SQ SEQUENCE 662 AA; 71877 MW; CF3C66350EC46E20 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 662;  
 Best Local Similarity 20.0%; Pred. No. 3.4;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
 Db 7 CGSAAASAC 16

RESULT 46

060E18 ORYSA

ID 060E18\_ORYSA PRELIMINARY; PRT; 691 AA.

AC 060E18\_1

DT 25-OCT-2004 (T-EMBLrel. 28, Created)

DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)

DE Putative cyelic nucleotide gated ion channel.

GN Name=OSUNBA0017K09.10;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzae; Oryza.

OX NCBI\_TaxID=39947;

RN (1)

RP NUCLEOTIDE SEQUENCE.

RA Chow T.-Y., Heing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,

RA Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,

RA Chao Y.-L., Cheng C.-H., Chung C.-H., Han S.-Y., Hsiao S.-H.,

RA Heing Y.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,

RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,

Wu H.-P., Shaw J.-F.

RT "Oryza sativa BAC OSUNBA0017K09 genomic sequence."

RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC130597; AAU90233.1; -; Genomic DNA.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005216; P:ion channel activity; IEA.

DR GO; GO:0006811; P:ion transport; IEA.

DR InterPro; IPR005953; cNMP\_binding.

DR InterPro; IPR005821; Ion\_trans.

DR Pfam; PF00027; cNMP\_binding; 1.

DR Pfam; PF00520; Ion\_trans; 1.

DR SMART; SM00100; cNMP; 1.

DR PROSITE; PS50042; cNMP\_BINDING\_3; 1.

SQ SEQUENCE 691 AA; 77220 MW; D059333316F2E299 CRC64;

Query Match 65.4%; Score 17; DB 2; Length 691;  
 Best Local Similarity 20.0%; Pred. No. 3.5;  
 Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10  
 Db 298 CASSTAACVAC 307

RESULT 47

0819P8 9BIVA

ID 0819P8\_9BIVA PRELIMINARY; PRT; 699 AA.

AC 0819P8\_1

DT 01-MAR-2003 (T-EMBLrel. 23, Created)

DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)

DE Alpha-amylose.

OS Name=amy;

GN Corbicula fluminea.

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;

OC Corbiculoidea; Corbiculidae; Corbicula.

OX NCBI\_TaxID=45949;

RN (1)

RP NUCLEOTIDE SEQUENCE.

RA Da Lage J.-L., Van Wormhout A., Carlot M.-L.;

RT "Diversity and evolution of the alpha-amylose genes in Animals."

RL Biologia 57:181-189(2002).

RN (2)

RP NUCLEOTIDE SEQUENCE.

RA Published=14704857; DOI=10.1007/s00018-003-3334-Y;

RT Da Lage J.-L., Feller G., Janacek S.;

RT "Horizontal gene transfer from Eukarya to bacteria and domain

shuffling: the alpha-amylose model."

RT Cell. Mol. Life Sci. 61:97-109(2004).

RN (3)

RP NUCLEOTIDE SEQUENCE.

RA Da Lage J.-L.;

RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF468016; ANO17927.2; -; Genomic DNA.

DR HSSP; P04745; 1SMD.

DR GO; GO:0004556; P:alpha-amylose activity; IEA.

DR GO; GO:0005509; P:calcium ion binding; IEA.

DR GO; GO:0016798; P:hydrolase activity, acting on glycosyl bonds; IEA.

DR GO; GO:0005975; P:carbohydrate metabolism; IEA.

DR InterPro; IPR006048; Alpha\_amy1\_C.

DR InterPro; IPR006047; Alpha\_amy1\_cat.

DR InterPro; IPR006589; Alp\_amy1\_cat\_sub.

DR InterPro; IPR006046; Glyco\_hydro\_13.

DR Pfam; PF00128; Alpha-amylose; 1.

DR Pfam; PF02806; Alpha-amylose\_C; 1.

```

DR PRINTS: PR00110; ALPHAMYLASE.
DR SMART; SM00642; Amy; 1.
DR SMART; SM00632; Amy; C; 1.
SQ SEQUENCE 699 AA; 76544 MW; 25D57008165B04CB CRC64;

Query Match
Best Local Similarity 20.0%; Pred. No. 3.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 544 CTSDATSAC 553

RESULT 48
Q5L6X2_CHLAB
ID Q5L6X2_CHLAB PRELIMINARY; PRT; 705 AA.
AC Q5L6X2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Putative transport protein.
GN OrderedLocusNames=CAB140;
OS Chlamydomophila abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83555;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=526/3;
RX PubMed=15837807; DOI=10.1101/gr.3684805;
RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D.,
RA Livingstone M., Cerdano-Tarraga A.-M., Harris B., Doggett J.,
RA Ormond D., Mungall K., Clarke K., Felwell T., Hance Z., Sanders M.,
RA Quail M.A., Price C., Barrett B.G., Parkhill J., Longbottom D.,
RT "The Chlamydomophila abortus genome sequence reveals an array of
RT variable proteins that contribute to interspecies variation.";
RL Genome Res. 15:629-640 (2005).
DR EMBL; CR848038; CAH63598.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 705 AA; 81369 MW; 1997374B5266E2AB CRC64;

Query Match
Best Local Similarity 20.0%; Pred. No. 3.5;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 67 CSSLSATTC 76

RESULT 49
Q9NKE3_DROME
ID Q9NKE3_DROME PRELIMINARY; PRT; 734 AA.
AC Q9NKE3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Hypothetical protein BG:DS00180.5.
GN Name=BG:DS00180.5; ORFNames=CG31765, CG31766, CG31841;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RX MEDLINE=99403001; PubMed=10471707;
RA Ashburner M., Mitra S., Roote J., Lewis S.E., Blazej R.G., Davis T.,
RA Doyle C., Galle R.F., George R.A., Harris N.L., Hartzell G.,
RA Harvey D.A., Hong L., Houston K.A., Hoskins R.A., Johnson G.,
RA Martin C., Moshrefi A.R., Palazzolo M., Reese M.G., Spradling A.C.,
RA Teang G., Wan K.H., Whitelaw K., Celniker S.E., Rubin G.M.;

```

```

RT "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the adh region.";
RL Genetics 153:179-219(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkeley;
RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter B., Blazej R.G.,
RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Fartan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummachi S.R., Kaira K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomolan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacle J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Snit E., Swirekas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003408; AAF44838.1; -; Genomic_DNA.
DR FlyBase; FBgn0051765; CG31765.
DR FlyBase; FBgn0051766; CG31766.
DR FlyBase; FBgn0051841; CG31841.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004245; P:peptidyl-dipeptidase A activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001548; Peptidase_M2.
DR Pfam; PF01401; Peptidase_M2; 1.
KW Hypothetical protein.
SQ SEQUENCE 734 AA; 79831 MW; 4C0B59393AF9FC1E CRC64;

Query Match
Best Local Similarity 20.0%; Pred. No. 3.6;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 681 CSSASAVASC 690

RESULT 50
Q4YNR2_PLABE
ID Q4YNR2_PLABE PRELIMINARY; PRT; 784 AA.
AC Q4YNR2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=PB000881.03.0;
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5821;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karras M., Raine J.D., Carlton J.M., Koof T.W.A.,
RA Britman M., Florens L., Jansen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churche C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrett B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic,
RT transcriptomic, and proteomic analyses.";
RL Science 307:82-86(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CA01003291; CA100348.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 784 AA; 91429 MW; B76DA6BAD7A0C1D6 CRC64;

Query Match
Best Local Similarity 20.0%; Pred. No. 3.7;
Matches 2; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 10
DB 363 CSSSFFSSSC 372

```

Thu Jan 5 09:14:13 2006

us-09-932-322-11.rup

Page 25

Search completed: January 4, 2006, 16:09:55  
Job time : 67.3043 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 4, 2006, 14:57:44 : Search time 82.1217 Seconds  
(without alignments)  
64.204 Million cell updates/sec

Title: US-09-932-322-12  
Perfect score: 28  
Sequence: 1 CXXXXXXXXXXC 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : A\_Geneseq.21.\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*  
9: geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	64.3	104	7	ABO82562 Pseudomon
2	18	64.3	110	8	ADP30723 Human sec
3	18	64.3	135	8	ADP31638 Human sec
4	18	64.3	142	7	ABO79745 Pseudomon
5	18	64.3	178	7	ABO73284 Pseudomon
6	18	64.3	264	8	ADP31527 Human sec
7	18	64.3	270	8	ADP31435 Human sec
8	18	64.3	307	4	ABH65879 Drosophil
9	18	64.3	357	8	ADP31223 Human sec
10	18	64.3	615	8	ADP31132 Human sec
11	18	64.3	627	6	AAU51580 Propionib
12	18	64.3	627	6	ABM48099 Pseudomon
13	18	64.3	669	8	ADP31142 Human sec
14	18	64.3	711	8	ADP31215 Human sec
15	18	64.3	739	8	ADP31196 Human sec
16	18	64.3	771	8	ADP31244 Human sec
17	18	64.3	876	8	ADP31220 Human sec
18	18	64.3	882	8	ADP31688 Human sec
19	18	64.3	925	5	AAO14246 Human pre
20	18	64.3	1044	8	ADP31517 Human sec
21	18	64.3	1086	8	ADP31175 Human sec
22	18	64.3	1113	8	ADP31508 Human sec
23	18	64.3	1128	6	ADA15725 C. elegan
24	18	64.3	1134	8	ADP30741 Human sec

25	18	64.3	1134	8	ADP30924 Human sec
26	18	64.3	1168	8	ADP31046 Human sec
27	18	64.3	1191	8	ADP30993 Human sec
28	18	64.3	1260	8	ADP31533 Human sec
29	18	64.3	1289	8	ADP30675 Human sec
30	18	64.3	1337	8	ADP31357 Human sec
31	18	64.3	1454	8	ADP31177 Human sec
32	18	64.3	1480	8	ADP30557 Human sec
33	18	64.3	1588	5	ABH09437 H. Influe
34	18	64.3	1617	8	ADP30660 Human sec
35	18	64.3	1652	6	ADA15715 C. elegan
36	18	64.3	1725	8	ADP30654 Human sec
37	18	64.3	1833	8	ADP30642 Human sec
38	18	64.3	2088	8	ADP31178 Human sec
39	18	64.3	2127	8	ADP31327 Human sec
40	18	64.3	2484	8	ADP66690 Human mis
41	18	64.3	2508	8	ADA15721 C. elegan
42	18	64.3	2544	6	ADA15717 C. elegan
43	18	64.3	2601	6	ADA15723 C. elegan
44	18	64.3	2833	8	ADP31299 Human sec
45	18	64.3	2835	8	ADP30572 Human sec
46	18	64.3	3411	8	ADP30667 Human sec
47	18	64.3	4440	6	ABH88256 Human sec
48	18	64.3	4440	6	ABU90135 Human sec
49	18	64.3	4440	6	ABU96437 Human sec
50	18	64.3	4440	6	ABU99046 Human sec
51	18	64.3	4440	6	ABU98261 Human sec
52	18	64.3	4440	6	ABU91967 Human sec
53	18	64.3	4440	6	ABU85271 Human sec
54	18	64.3	4440	6	ABO00410 Human sec
55	18	64.3	4440	6	ABU88961 Human sec
56	18	64.3	4440	6	ABO06457 Human sec
57	18	64.3	4440	6	ABU95517 Human sec
58	18	64.3	4440	6	ABU95207 Human sec
59	18	64.3	4440	6	ABU90755 Human sec
60	18	64.3	4440	6	ABU93917 Human sec
61	18	64.3	4440	6	ABU86191 Human sec
62	18	64.3	4440	6	ABU82046 Human sec
63	18	64.3	4440	6	ABU07907 Human sec
64	18	64.3	4440	6	ABU94227 Human sec
65	18	64.3	4440	6	ABO00100 Human sec
66	18	64.3	4440	6	ABU87111 Human sec
67	18	64.3	4440	6	ABU91352 Human sec
68	18	64.3	4440	6	ABU90445 Human sec
69	18	64.3	4440	6	ABU97036 Human sec
70	18	64.3	4440	6	ABO05232 Human sec
71	18	64.3	8976	8	ADP31425 Human sec
72	18	64.3	9195	8	ADP31494 Human sec
73	17	60.7	31	8	ADP30526 Human sec
74	17	60.7	31	8	ADP30526 Human sec
75	17	60.7	53	3	AAU58907 Arabidops
76	17	60.7	65	4	ABG29760 Arabidops
77	17	60.7	99	8	ADP30709 Human sec
78	17	60.7	105	7	ABO78139 Pseudomon
79	17	60.7	119	7	ABO79679 Pseudomon
80	17	60.7	149	3	AAU94985 Human sec
81	17	60.7	160	8	ADP31048 Human sec
82	17	60.7	165	8	ADP31515 Human sec
83	17	60.7	165	8	ADP31516 Human sec
84	17	60.7	168	7	ABO79318 Human sec
85	17	60.7	168	8	ADP31621 Human sec
86	17	60.7	168	8	ADP31099 Human sec
87	17	60.7	181	8	ADP32231 Plant ful
88	17	60.7	234	7	ABO68370 Pseudomon
89	17	60.7	240	8	ADP30737 Human sec
90	17	60.7	252	8	ADP31487 Human sec
91	17	60.7	264	7	ABO75516 Pseudomon
92	17	60.7	264	8	ADP31272 Human sec
93	17	60.7	297	8	ADP31312 Human sec
94	17	60.7	309	8	ADP30862 Human sec
95	17	60.7	309	8	ADP30873 Human sec
96	17	60.7	320	8	ADP31607 Human sec
97	17	60.7	320	8	ADP31649 Human sec

98	17	60.7	325	7	ABM87858	Abm87858	Rice	ab10	171	17	60.7	950	8	ADP31167	ADP31167	Human	sec
99	17	60.7	330	7	ADB70062	ADB70062	C. neofo	r	172	17	60.7	951	8	ADP30943	ADP30943	Human	sec
100	17	60.7	348	8	ADP31441	ADP31441	Human	sec	173	17	60.7	957	8	ADP31528	ADP31528	Human	sec
101	17	60.7	357	8	ADP31267	ADP31267	Human	sec	174	17	60.7	957	8	ADT07508	ADT07508	Human	col
102	17	60.7	367	8	ADP30950	ADP30950	Human	sec	175	17	60.7	960	8	ADP31471	ADP31471	Human	sec
103	17	60.7	381	8	ADP30655	ADP30655	Human	sec	176	17	60.7	960	8	ADP31470	ADP31470	Human	sec
104	17	60.7	384	8	ADP30656	ADP30656	Human	sec	177	17	60.7	981	8	ADP30547	ADP30547	Human	sec
105	17	60.7	392	8	ADP31061	ADP31061	Human	sec	178	17	60.7	996	8	ADP31538	ADP31538	Human	sec
106	17	60.7	392	8	ADP31055	ADP31055	Human	sec	179	17	60.7	1035	8	ADP31552	ADP31552	Human	sec
107	17	60.7	393	8	ADP31345	ADP31345	Human	sec	180	17	60.7	1038	8	ADP30860	ADP30860	Human	sec
108	17	60.7	402	8	ADP30742	ADP30742	Human	sec	181	17	60.7	1044	8	ADP31550	ADP31550	Human	sec
109	17	60.7	402	8	ADP31414	ADP31414	Human	sec	182	17	60.7	1048	8	ADP31642	ADP31642	Human	sec
110	17	60.7	402	8	ADP31262	ADP31262	Human	sec	183	17	60.7	1053	8	ADP30886	ADP30886	Human	sec
111	17	60.7	409	7	ABO79810	ABO79810	Pseudomon		184	17	60.7	1059	8	ADP31042	ADP31042	Human	sec
112	17	60.7	411	8	ADP30729	ADP30729	Human	sec	185	17	60.7	1065	8	ADP31347	ADP31347	Human	sec
113	17	60.7	412	8	ADT57242	ADT57242	Plant	pol	186	17	60.7	1076	6	ABU81145	ABU81145	Human	PRO
114	17	60.7	420	8	ADP31273	ADP31273	Human	sec	187	17	60.7	1076	6	ABU66845	ABU66845	Human	PRO
115	17	60.7	420	8	ADP31274	ADP31274	Human	sec	188	17	60.7	1082	8	ADP30934	ADP30934	Human	sec
116	17	60.7	423	8	ADP31479	ADP31479	Human	sec	189	17	60.7	1092	8	ADP31358	ADP31358	Human	sec
117	17	60.7	483	8	ADP30852	ADP30852	Human	sec	190	17	60.7	1101	8	ADP31462	ADP31462	Human	sec
118	17	60.7	483	8	ADP30853	ADP30853	Human	sec	191	17	60.7	1110	8	ADP31430	ADP31430	Human	sec
119	17	60.7	510	8	ADP31049	ADP31049	Human	sec	192	17	60.7	1119	8	ADP31452	ADP31452	Human	sec
120	17	60.7	525	8	ADP31227	ADP31227	Human	sec	193	17	60.7	1128	8	ADP30483	ADP30483	Human	sec
121	17	60.7	528	8	ADP30529	ADP30529	Human	sec	194	17	60.7	1134	8	ADP30647	ADP30647	Human	sec
122	17	60.7	531	8	ADP31539	ADP31539	Human	sec	195	17	60.7	1134	8	ADP31537	ADP31537	Human	sec
123	17	60.7	531	8	ADP31540	ADP31540	Human	sec	196	17	60.7	1141	8	ADP30677	ADP30677	Human	sec
124	17	60.7	543	8	ADP30864	ADP30864	Human	sec	197	17	60.7	1170	8	ADP30819	ADP30819	Human	sec
125	17	60.7	552	8	ADP31019	ADP31019	Human	sec	198	17	60.7	1170	8	ADP30922	ADP30922	Human	sec
126	17	60.7	555	8	ADP31168	ADP31168	Human	sec	199	17	60.7	1173	8	ADP31155	ADP31155	Human	sec
127	17	60.7	559	8	ADP30716	ADP30716	Human	sec	200	17	60.7	1179	8	ADP30486	ADP30486	Human	sec
128	17	60.7	564	8	ADP30743	ADP30743	Human	sec	201	17	60.7	1189	8	ADP31043	ADP31043	Human	sec
129	17	60.7	567	8	ADP31565	ADP31565	Human	sec	202	17	60.7	1197	8	ADP31342	ADP31342	Human	sec
130	17	60.7	574	8	ADP31133	ADP31133	Human	sec	203	17	60.7	1205	8	ADP30935	ADP30935	Human	sec
131	17	60.7	592	8	ADP30917	ADP30917	Human	sec	204	17	60.7	1252	8	ADP30678	ADP30678	Human	sec
132	17	60.7	624	8	ADP31324	ADP31324	Human	sec	205	17	60.7	1358	8	ADP30995	ADP30995	Human	sec
133	17	60.7	624	8	ADP31325	ADP31325	Human	sec	206	17	60.7	1362	8	ADP31181	ADP31181	Human	sec
134	17	60.7	629	8	ADP31295	ADP31295	Human	sec	207	17	60.7	1383	8	ADP31091	ADP31091	Human	sec
135	17	60.7	645	8	ADP30858	ADP30858	Human	sec	208	17	60.7	1417	8	ADP31160	ADP31160	Human	sec
136	17	60.7	666	8	ADP31256	ADP31256	Human	sec	209	17	60.7	1420	8	ADP30944	ADP30944	Human	sec
137	17	60.7	681	8	ADP30598	ADP30598	Human	sec	210	17	60.7	1435	6	ABU09754	ABU09754	Human	sec
138	17	60.7	702	8	ADP30868	ADP30868	Human	sec	211	17	60.7	1435	6	ABU90136	ABU90136	Human	sec
139	17	60.7	702	8	ADP31518	ADP31518	Human	sec	212	17	60.7	1435	6	ABU96438	ABU96438	Human	sec
140	17	60.7	702	8	ADP31277	ADP31277	Human	sec	213	17	60.7	1435	6	ABU99047	ABU99047	Human	sec
141	17	60.7	702	8	ADP31278	ADP31278	Human	sec	214	17	60.7	1435	6	ABU98262	ABU98262	Human	sec
142	17	60.7	708	8	ADP31191	ADP31191	Human	sec	215	17	60.7	1435	6	ABU91968	ABU91968	Human	sec
143	17	60.7	711	8	ADP31535	ADP31535	Human	sec	216	17	60.7	1435	6	ABU85272	ABU85272	Human	sec
144	17	60.7	714	8	ADP31440	ADP31440	Human	sec	217	17	60.7	1435	6	ABU00411	ABU00411	Human	sec
145	17	60.7	714	8	ADP31561	ADP31561	Human	sec	218	17	60.7	1435	6	ABU88962	ABU88962	Human	sec
146	17	60.7	759	8	ADP31141	ADP31141	Human	sec	219	17	60.7	1435	6	ABU06458	ABU06458	Human	sec
147	17	60.7	768	8	ADP30912	ADP30912	Human	sec	220	17	60.7	1435	6	ABU95518	ABU95518	Human	sec
148	17	60.7	771	8	ADP30907	ADP30907	Human	sec	221	17	60.7	1435	6	ABU95208	ABU95208	Human	sec
149	17	60.7	774	8	ADP30885	ADP30885	Human	sec	222	17	60.7	1435	6	ABU90756	ABU90756	Human	sec
150	17	60.7	795	8	ADP31332	ADP31332	Human	sec	223	17	60.7	1435	6	ABU93918	ABU93918	Human	sec
151	17	60.7	799	8	ADP31261	ADP31261	Human	sec	224	17	60.7	1435	6	ABU86192	ABU86192	Human	sec
152	17	60.7	813	8	ADP30649	ADP30649	Human	sec	225	17	60.7	1435	6	ABU82047	ABU82047	Human	sec
153	17	60.7	834	8	ADP30738	ADP30738	Human	sec	226	17	60.7	1435	6	ABU07908	ABU07908	Human	sec
154	17	60.7	843	8	ADP30663	ADP30663	Human	sec	227	17	60.7	1435	6	ABU94428	ABU94428	Human	sec
155	17	60.7	849	8	ADP31037	ADP31037	Human	sec	228	17	60.7	1435	6	ABU00101	ABU00101	Human	sec
156	17	60.7	849	8	ADP31182	ADP31182	Human	sec	229	17	60.7	1435	6	ABU87112	ABU87112	Human	sec
157	17	60.7	849	8	ADP31275	ADP31275	Human	sec	230	17	60.7	1435	6	ABU91353	ABU91353	Human	sec
158	17	60.7	849	8	ADP31276	ADP31276	Human	sec	231	17	60.7	1435	6	ABU90446	ABU90446	Human	sec
159	17	60.7	861	8	ADP31340	ADP31340	Human	sec	232	17	60.7	1435	6	ABU97037	ABU97037	Human	sec
160	17	60.7	882	8	ADP30487	ADP30487	Human	sec	233	17	60.7	1435	6	ABU05233	ABU05233	Human	sec
161	17	60.7	895	8	ADU92573	ADU92573	Human	PRO	234	17	60.7	1464	8	ADP31437	ADP31437	Human	sec
162	17	60.7	900	6	ADP31337	ADP31337	Human	sec	235	17	60.7	1472	8	ADP31611	ADP31611	Human	sec
163	17	60.7	908	6	ABU08492	ABU08492	Alpha-he		236	17	60.7	1530	8	ADP31536	ADP31536	Human	sec
164	17	60.7	921	8	ADP31582	ADP31582	Human	sec	237	17	60.7	1550	8	ADP30567	ADP30567	Human	sec
165	17	60.7	930	8	ADP31444	ADP31444	Human	sec	238	17	60.7	1560	8	ADP31631	ADP31631	Human	sec
166	17	60.7	933	8	ADP31140	ADP31140	Human	sec	239	17	60.7	1614	8	ADP31529	ADP31529	Human	sec
167	17	60.7	939	8	ADP31542	ADP31542	Human	sec	240	17	60.7	1647	8	ADP30670	ADP30670	Human	sec
168	17	60.7	939	8	ADP31086	ADP31086	Human	sec	241	17	60.7	1662	8	ADP31419	ADP31419	Human	sec
169	17	60.7	939	8	ADP31541	ADP31541	Human	sec	242	17	60.7	1662	8	ADP31513	ADP31513	Human	sec
170	17	60.7	947	8	ADP30937	ADP30937	Human	sec	243	17	60.7	1700	3	AA18144	AA18144	Plaemodiu	

244	17	60.7	1743	6	ABU88255	Abu88255	Novel	hum	317	17	60.7	3390	8	ADP31148	Adp31148	Human	sec
245	17	60.7	1743	6	ABU90134	Abu90134	Novel	hum	318	17	60.7	3447	8	ADP31112	Adp31112	Human	sec
246	17	60.7	1743	6	ABU9436	Abu9436	Novel	hum	319	17	60.7	4315	5	ADP43908	Adp43908	MUC5B	par
247	17	60.7	1743	6	ABU99045	Abu99045	Novel	hum	320	17	60.7	5304	8	ADP30706	Adp30706	Human	sec
248	17	60.7	1743	6	ABU98260	Abu98260	Novel	hum	321	17	60.7	5514	8	ADP31186	Adp31186	Human	sec
249	17	60.7	1743	6	ABU91966	Abu91966	Novel	hum	322	17	60.7	5514	8	ADP31591	Adp31591	Human	sec
250	17	60.7	1743	6	ABU85270	Abu85270	Novel	hum	323	17	60.7	5703	8	ADL23265	Adl23265	Human	MUC
251	17	60.7	1743	6	ABO00409	AbO00409	Novel	hum	324	17	60.7	5858	9	ABE56507	Aeb56507	Radiochem	
252	17	60.7	1743	6	ABU88960	Abu88960	Novel	hum	325	17	60.7	6465	8	ADP30705	Adp30705	Human	sec
253	17	60.7	1743	6	ABO06456	AbO06456	Novel	hum	326	17	60.7	6729	8	ADP31600	Adp31600	Human	sec
254	17	60.7	1743	6	ABU95516	Abu95516	Novel	hum	327	17	60.7	7285	6	ABJ38280	Abj38280	pMG21-RA	
255	17	60.7	1743	6	ABU95206	Abu95206	Novel	hum	328	17	60.7	7339	8	AAO16358	Aao16358	Human	tra
256	17	60.7	1743	6	ABU90754	Abu90754	Novel	hum	329	17	60.7	10944	6	ADP31311	Adp31311	Human	sec
257	17	60.7	1743	6	ABU93916	Abu93916	Novel	hum	330	17	60.7	11328	8	ADP31310	Adp31310	Human	sec
258	17	60.7	1743	6	ABU86190	Abu86190	Novel	hum	331	16	57.1	42	9	ADZ47464	Aez47464	Optum	pop
259	17	60.7	1743	6	ABU82045	Abu82045	Novel	hum	332	16	57.1	46	9	AEA34334	Aea34334	Pig	EST
260	17	60.7	1743	6	ABU07906	Abu07906	Novel	hum	333	16	57.1	49	4	AAW24449	Aaw24449	Pig	EST
261	17	60.7	1743	6	ABU94226	Abu94226	Novel	hum	334	16	57.1	51	4	AAU57122	Aau57122	Proprionib	
262	17	60.7	1743	6	ABO00099	AbO00099	Novel	hum	335	16	57.1	51	6	ABM53641	Abm53641	Proprionib	
263	17	60.7	1743	6	ABU87110	Abu87110	Novel	hum	336	16	57.1	52	4	AAU59153	Aau59153	Proprionib	
264	17	60.7	1743	6	ABU91351	Abu91351	Novel	hum	337	16	57.1	52	6	ABM55672	Abm55672	Proprionib	
265	17	60.7	1743	6	ABU90444	Abu90444	Novel	hum	338	16	57.1	58	9	ADZ11761	Adz11761	E. coli	D
266	17	60.7	1743	6	ABU97035	Abu97035	Novel	hum	339	16	57.1	85	4	AAU61946	Aau61946	Proprionib	
267	17	60.7	1743	6	ABO05231	AbO05231	Novel	hum	340	16	57.1	85	6	ABM58465	Abm58465	Proprionib	
268	17	60.7	1776	9	ADY62715	Ady62715	Human	alp	341	16	57.1	94	4	AAAB65963	Aab65963	Human	sec
269	17	60.7	1782	8	ADP31270	Adp31270	Human	sec	342	16	57.1	100	8	ADKX5826	Adkx5826	Plant	ful
270	17	60.7	1803	8	ADP30562	Adp30562	Human	sec	343	16	57.1	109	8	ADP30690	Adp30690	Human	sec
271	17	60.7	1894	6	ABU88253	Abu88253	Novel	hum	344	16	57.1	132	8	ADQ94844	Adq94844	Generic	b
272	17	60.7	1894	6	ABU90132	Abu90132	Novel	hum	345	16	57.1	133	7	ABO70871	AbO70871	Pseudomon	
273	17	60.7	1894	6	ABU90132	Abu90132	Novel	hum	346	16	57.1	134	8	ADQ94811	Adq94811	Generic	b
274	17	60.7	1894	6	ABU96434	Abu96434	Novel	hum	347	16	57.1	138	7	ABO67837	AbO67837	Pseudomon	
275	17	60.7	1894	6	ABU99043	Abu99043	Novel	hum	348	16	57.1	142	7	ABO75180	AbO75180	Pseudomon	
276	17	60.7	1894	6	ABU98258	Abu98258	Novel	hum	349	16	57.1	144	7	ABO77101	AbO77101	Pseudomon	
277	17	60.7	1894	6	ABU91964	Abu91964	Novel	hum	350	16	57.1	144	8	ADX72498	Adx72498	Plant	ful
278	17	60.7	1894	6	ABU85258	Abu85258	Novel	hum	351	16	57.1	151	7	ABO81297	AbO81297	Pseudomon	
279	17	60.7	1894	6	ABO00407	AbO00407	Novel	hum	352	16	57.1	154	7	ABO80026	AbO80026	Pseudomon	
280	17	60.7	1894	6	ABO08958	AbO08958	Novel	hum	353	16	57.1	155	7	ABO71527	AbO71527	Pseudomon	
281	17	60.7	1894	6	ABO06454	AbO06454	Novel	hum	354	16	57.1	159	8	ADP30763	Adp30763	Human	sec
282	17	60.7	1894	6	ABU95514	Abu95514	Novel	hum	355	16	57.1	159	8	ADP30762	Adp30762	Human	sec
283	17	60.7	1894	6	ABU95204	Abu95204	Novel	hum	356	16	57.1	160	4	AAW95024	Aaw95024	Human	rep
284	17	60.7	1894	6	ABU90752	Abu90752	Novel	hum	357	16	57.1	160	4	ABB95732	Abb95732	Human	tes
285	17	60.7	1894	6	ABU93914	Abu93914	Novel	hum	358	16	57.1	165	8	ADP31286	Adp31286	Human	sec
286	17	60.7	1894	6	ABU86188	Abu86188	Novel	hum	359	16	57.1	165	8	ADP31174	Adp31174	Human	sec
287	17	60.7	1894	6	ABU82043	Abu82043	Novel	hum	360	16	57.1	171	8	ADP30794	Adp30794	Human	sec
288	17	60.7	1894	6	ABU07904	Abu07904	Novel	hum	361	16	57.1	171	8	ADP31497	Adp31497	Human	sec
289	17	60.7	1894	6	ABU94224	Abu94224	Novel	hum	362	16	57.1	171	8	ADP30793	Adp30793	Human	sec
290	17	60.7	1894	6	ABO00097	AbO00097	Novel	hum	363	16	57.1	180	8	ADP30820	Adp30820	Human	sec
291	17	60.7	1894	6	ABU87108	Abu87108	Novel	hum	364	16	57.1	180	8	ADP30825	Adp30825	Human	sec
292	17	60.7	1894	6	ABU91349	Abu91349	Novel	hum	365	16	57.1	180	8	ADP30828	Adp30828	Human	sec
293	17	60.7	1894	6	ABU90442	Abu90442	Novel	hum	366	16	57.1	180	8	ADP30821	Adp30821	Human	sec
294	17	60.7	1894	6	ABU97033	Abu97033	Novel	hum	367	16	57.1	180	8	ADP30826	Adp30826	Human	sec
295	17	60.7	1894	6	ABO05229	AbO05229	Novel	hum	368	16	57.1	180	8	ADP30827	Adp30827	Human	sec
296	17	60.7	1968	8	ADP30689	Adp30689	Human	sec	369	16	57.1	188	8	ADP31660	Adp31660	Human	sec
297	17	60.7	2058	8	ADP31630	Adp31630	Human	sec	370	16	57.1	195	8	ADP30590	Adp30590	Human	sec
298	17	60.7	2091	8	ADP31088	Adp31088	Human	sec	371	16	57.1	198	2	AAE59841	Aae59841	Agora4L	pr
299	17	60.7	2109	8	ADT07505	AdT07505	Human	col	372	16	57.1	198	8	AAE92113	Aae92113	Human	ApO
300	17	60.7	2124	8	ADP30560	Adp30560	Human	sec	373	16	57.1	198	8	ADP30766	Adp30766	Human	sec
301	17	60.7	2187	8	ADP30644	Adp30644	Human	sec	374	16	57.1	198	8	ADP30767	Adp30767	Human	sec
302	17	60.7	2254	8	ADP30862	Adp30862	Human	sec	375	16	57.1	199	7	ABO73562	AbO73562	Pseudomon	
303	17	60.7	2260	8	ADT07506	AdT07506	Human	col	376	16	57.1	201	8	ADP30711	Adp30711	Human	sec
304	17	60.7	2272	8	ADP30687	Adp30687	Human	sec	377	16	57.1	215	9	ABE803544	Aeb803544	Mycobacte	
305	17	60.7	2272	8	ADP31136	Adp31136	Human	sec	378	16	57.1	215	9	AEA79447	Aea79447	Novel	M.
306	17	60.7	2304	8	ADP30669	Adp30669	Human	sec	379	16	57.1	217	7	ABO72885	AbO72885	Pseudomon	
307	17	60.7	2307	8	ADP31252	Adp31252	Human	sec	380	16	57.1	219	7	ABO76425	AbO76425	Pseudomon	
308	17	60.7	2307	8	ADP31394	Adp31394	Human	sec	381	16	57.1	225	8	ADP30589	Adp30589	Human	sec
309	17	60.7	2349	8	ADP30959	Adp30959	Human	sec	382	16	57.1	228	8	ADP31281	Adp31281	Human	sec
310	17	60.7	2382	8	ADP31341	Adp31341	Human	sec	383	16	57.1	229	3	ADCO7930	AdcO7930	Rice	prot
311	17	60.7	2401	8	ADT07507	AdT07507	Human	col	384	16	57.1	234	3	AAE23898	Aae23898	Arabidops	
312	17	60.7	2415	8	ADP31023	Adp31023	Human	sec	385	16	57.1	234	5	AAE49556	Aae49556	Arabidops	
313	17	60.7	2418	8	ADP31105	Adp31105	Human	sec	386	16	57.1	234	5	AAE93028	Aae93028	Arabidops	
314	17	60.7	2616	8	ADP31253	Adp31253	Human	sec	387	16	57.1	234	7	ADD30781	AdD30781	Plant	yle
315	17	60.7	2976	9	AEA96675	Aeb96675	N. mening		388	16	57.1	234	7	ADL43959	Adl43959	Plant	tra
316	17	60.7	3339	8	ADP30724	Adp30724	Human	sec	389	16	57.1	234	8	ADL61385	Adl61385	A. thalia	

390	16	57.1	234	8	AD002129	Ad002129	Thalactres	463	16	57.1	471	8	ADP30854	ADp30854	Human sec
391	16	57.1	234	8	ADP30484	Adp30484	Human sec	464	16	57.1	472	8	ADP31222	ADp31222	Human sec
392	16	57.1	234	8	ADV09549	Adv09549	A. thalia	465	16	57.1	473	8	ADP31083	ADp31083	Human sec
393	16	57.1	234	8	ADV09550	Adv09550	A. thalia	466	16	57.1	474	8	ADP31235	ADp31235	Human sec
394	16	57.1	243	8	ADP31343	Adp31343	Human sec	467	16	57.1	480	8	ADP31484	ADp31484	Human sec
395	16	57.1	246	8	ADP30619	Adp30619	Human sec	468	16	57.1	489	8	ADP31380	ADp31380	Human sec
396	16	57.1	249	7	AB069987	Ab069987	Pseudomon	469	16	57.1	493	9	ADW69641	ADw69641	HIV-1 GP1
397	16	57.1	254	8	ADP30739	Adp30739	Human sec	470	16	57.1	501	8	ADP31689	ADp31689	Human sec
398	16	57.1	254	8	ADP30693	Adp30693	Human sec	471	16	57.1	510	8	ADP31549	ADp31549	Human sec
399	16	57.1	278	8	ADQ66574	Adq66574	Novel hum	472	16	57.1	519	8	ADP31190	ADp31190	Human sec
400	16	57.1	279	8	ADP31523	Adp31523	Human sec	473	16	57.1	528	8	ADP31525	ADp31525	Human sec
401	16	57.1	279	8	ADP31489	Adp31489	Human sec	474	16	57.1	531	8	ADP31696	ADp31696	Human sec
402	16	57.1	280	7	AB072037	Ab072037	Pseudomon	475	16	57.1	533	8	ADP31279	ADp31279	Human sec
403	16	57.1	285	7	AB072271	Ab072271	Pseudomon	476	16	57.1	543	8	ADP31329	ADp31329	Human sec
404	16	57.1	286	8	ADP09918	Adp09918	Human pro	477	16	57.1	546	8	ADP31449	ADp31449	Human sec
405	16	57.1	291	8	ADP31193	Adp31193	Human sec	478	16	57.1	549	8	ADP30855	ADp30855	Human sec
406	16	57.1	294	8	ADP31045	Adp31045	Human sec	479	16	57.1	549	8	ADP31639	ADp31639	Human sec
407	16	57.1	306	7	AB074423	Ab074423	Pseudomon	480	16	57.1	549	8	ADP31009	ADp31009	Human sec
408	16	57.1	311	2	AAR59843	Aar59843	ApocRfx2	481	16	57.1	552	8	ADP31524	ADp31524	Human sec
409	16	57.1	311	2	AAR92115	Aar92115	Human Apo	482	16	57.1	557	8	ADP31103	ADp31103	Human sec
410	16	57.1	315	8	ADP31685	Adp31685	Human sec	483	16	57.1	558	8	ADP31254	ADp31254	Human sec
411	16	57.1	317	8	ADP31615	Adp31615	Human sec	484	16	57.1	558	8	ADP31255	ADp31255	Human sec
412	16	57.1	318	8	ADP30824	Adp30824	Human sec	485	16	57.1	563	8	ADP30881	ADp30881	Human sec
413	16	57.1	318	8	ADP31135	Adp31135	Human sec	486	16	57.1	585	8	ADP31458	ADp31458	Human sec
414	16	57.1	320	7	AB070026	Ab070026	Pseudomon	487	16	57.1	600	8	ADP30865	ADp30865	Human sec
415	16	57.1	331	7	AB070026	Ab070026	Pseudomon	488	16	57.1	604	8	ADP30940	ADp30940	Human sec
416	16	57.1	333	8	ADP30848	Adp30848	Human sec	489	16	57.1	604	8	ADP30941	ADp30941	Human sec
417	16	57.1	336	8	ADP31320	Adp31320	Human sec	490	16	57.1	605	8	ADL12887	ADl12887	Human bre
418	16	57.1	345	8	ADP31016	Adp31016	Human sec	491	16	57.1	610	8	ADP31264	ADp31264	Human sec
419	16	57.1	345	8	ADP31355	Adp31355	Human sec	492	16	57.1	612	8	ADP31169	ADp31169	Human sec
420	16	57.1	346	8	ADP30985	Adp30985	Human sec	493	16	57.1	618	8	ADP31467	ADp31467	Human sec
421	16	57.1	348	8	ADP30985	Adp30985	Human sec	494	16	57.1	618	8	ADP31466	ADp31466	Human sec
422	16	57.1	348	8	ADP31676	Adp31676	Human sec	495	16	57.1	626	8	ADP30528	ADp30528	Human sec
423	16	57.1	354	8	ADP30783	Adp30783	Human sec	496	16	57.1	627	8	ADP31388	ADp31388	Human sec
424	16	57.1	354	8	ADP30707	Adp30707	Human sec	497	16	57.1	629	7	AB068031	ABo68031	Pseudomon
425	16	57.1	354	8	ADP30779	Adp30779	Human sec	498	16	57.1	637	8	ADP30539	ADp30539	Human sec
426	16	57.1	354	8	ADP30784	Adp30784	Human sec	499	16	57.1	639	8	ADP30861	ADp30861	Human sec
427	16	57.1	354	8	ADP30778	Adp30778	Human sec	500	16	57.1	642	8	ADP31265	ADp31265	Human sec
428	16	57.1	357	8	ADP31560	Adp31560	Human sec	501	16	57.1	642	8	ADP30748	ADp30748	Human sec
429	16	57.1	357	8	ADP30505	Adp30505	Human sec	502	16	57.1	642	8	ADP31161	ADp31161	Human sec
430	16	57.1	360	8	ADP31439	Adp31439	Human sec	503	16	57.1	645	8	ADP31124	ADp31124	Human sec
431	16	57.1	363	8	ADP30659	Adp30659	Human sec	504	16	57.1	645	8	ADP31125	ADp31125	Human sec
432	16	57.1	369	8	ADP30807	Adp30807	Human sec	505	16	57.1	645	8	ADP31183	ADp31183	Human sec
433	16	57.1	372	8	ADP31492	Adp31492	Human sec	506	16	57.1	651	8	ADP31022	ADp31022	Human sec
434	16	57.1	372	8	ADP31610	Adp31610	Human sec	507	16	57.1	658	8	ADP31226	ADp31226	Human sec
435	16	57.1	378	8	ADP31375	Adp31375	Human sec	508	16	57.1	660	8	ADP31606	ADp31606	Human sec
436	16	57.1	381	8	ADP30580	Adp30580	Human sec	509	16	57.1	665	8	ADP30571	ADp30571	Human sec
437	16	57.1	382	8	ADP31221	Adp31221	Human sec	510	16	57.1	666	8	ADP30867	ADp30867	Human sec
438	16	57.1	384	8	ADP31338	Adp31338	Human sec	511	16	57.1	667	8	ADP04463	ADp04463	Sea squit
439	16	57.1	387	8	ADP31368	Adp31368	Human sec	512	16	57.1	669	8	ADP31598	ADp31598	Human sec
440	16	57.1	387	8	ADP31377	Adp31377	Human sec	513	16	57.1	672	8	ADP31667	ADp31667	Human sec
441	16	57.1	392	8	ADT59578	Adt59578	Plant pol	514	16	57.1	672	8	ADP31666	ADp31666	Human sec
442	16	57.1	396	8	ADP31378	Adp31378	Human sec	515	16	57.1	678	8	ADP31585	ADp31585	Human sec
443	16	57.1	399	8	ADP31464	Adp31464	Human sec	516	16	57.1	678	8	ADP31283	ADp31283	Human sec
444	16	57.1	411	7	ABM87022	Abm87022	Rice abio	517	16	57.1	681	8	ADP31053	ADp31053	Human sec
445	16	57.1	411	8	ADP31104	Adp31104	Human sec	518	16	57.1	690	8	ADP30750	ADp30750	Human sec
446	16	57.1	414	8	ADP31477	Adp31477	Human sec	519	16	57.1	699	8	ADP30747	ADp30747	Human sec
447	16	57.1	420	8	ADP31349	Adp31349	Human sec	520	16	57.1	705	8	ADP31330	ADp31330	Human sec
448	16	57.1	421	8	ADP31159	Adp31159	Human sec	521	16	57.1	705	8	ADP31623	ADp31623	Human sec
449	16	57.1	426	8	ADP31331	Adp31331	Human sec	522	16	57.1	710	8	ADP30746	ADp30746	Human sec
450	16	57.1	426	8	ADP31495	Adp31495	Human sec	523	16	57.1	711	8	ADP31652	ADp31652	Human sec
451	16	57.1	429	8	ADP30570	Adp30570	Human sec	524	16	57.1	728	8	ADP30508	ADp30508	Human sec
452	16	57.1	429	8	ADP31207	Adp31207	Human sec	525	16	57.1	729	8	ADP30541	ADp30541	Human sec
453	16	57.1	435	8	ADP31102	Adp31102	Human sec	526	16	57.1	746	8	ADP30982	ADp30982	Human sec
454	16	57.1	438	8	ADP30818	Adp30818	Human sec	527	16	57.1	750	8	ADP30576	ADp30576	Human sec
455	16	57.1	441	8	ADP31206	Adp31206	Human sec	528	16	57.1	750	8	ADP30857	ADp30857	Human sec
456	16	57.1	444	8	ADP31640	Adp31640	Human sec	529	16	57.1	750	8	ADP31131	ADp31131	Human sec
457	16	57.1	449	8	ADQ21283	Adq21283	Human sof	530	16	57.1	750	8	ADP30856	ADp30856	Human sec
458	16	57.1	450	8	ADP31085	Adp31085	Human sec	531	16	57.1	764	8	ADP31614	ADp31614	Human sec
459	16	57.1	453	8	ADP31465	Adp31465	Human sec	532	16	57.1	768	8	ADP31126	ADp31126	Human sec
460	16	57.1	459	8	ADP30932	Adp30932	Human sec	533	16	57.1	772	8	ADP30936	ADp30936	Human sec
461	16	57.1	459	8	ADP30930	Adp30930	Human sec	534	16	57.1	774	8	ADP30506	ADp30506	Human sec
462	16	57.1	467	8	ADX91801	Adx91801	Plant ful	535	16	57.1	774	8	ADP31373	ADp31373	Human sec

536	16	57.1	774	8	ADP30780	Human sec	609	16	57.1	1221	8	ADP30884	Adp30884	Human sec
537	16	57.1	774	8	ADP30781	Human sec	610	16	57.1	1224	8	ADP31426	Adp31426	Human sec
538	16	57.1	774	8	ADP31235	Human sec	611	16	57.1	1227	8	ADP31602	Adp31602	Human sec
539	16	57.1	776	7	ADP71157	Human int	612	16	57.1	1227	8	ADP31210	Adp31210	Human sec
540	16	57.1	776	7	ABM62451	Tumour-ab	613	16	57.1	1233	8	ADP30523	Adp30523	Human sec
541	16	57.1	776	9	ADY16233	PRO polyP	614	16	57.1	1239	8	ADP31297	Adp31297	Human sec
542	16	57.1	783	8	ADP31284	Human sec	615	16	57.1	1260	8	ADP30883	Adp30883	Human sec
543	16	57.1	783	8	ADP31436	Human sec	616	16	57.1	1269	8	ADP31189	Adp31189	Human sec
544	16	57.1	792	9	AEA0901	Novel hum	617	16	57.1	1269	8	ADJ23456	Adp23456	Human sec
545	16	57.1	804	8	ADP31635	Human sec	618	16	57.1	1282	8	ADP31328	Adp31328	Human sec
546	16	57.1	807	8	ADP31036	Human sec	619	16	57.1	1302	8	ADP31695	Adp31695	Human sec
547	16	57.1	813	8	ADP31282	Human sec	620	16	57.1	1305	8	ADP31389	Adp31389	Human sec
548	16	57.1	813	8	ADP30561	Human sec	621	16	57.1	1314	8	ADP31197	Adp31197	Human sec
549	16	57.1	816	8	ADP31245	Human sec	622	16	57.1	1344	8	ADP31211	Adp31211	Human sec
550	16	57.1	816	8	ADP31581	Human sec	623	16	57.1	1350	8	ADP31138	Adp31138	Human sec
551	16	57.1	828	8	ADP30878	Human sec	624	16	57.1	1359	8	ADP31129	Adp31129	Human sec
552	16	57.1	831	8	ADP31179	Human sec	625	16	57.1	1365	8	ADP31661	Adp31661	Human sec
553	16	57.1	852	8	ADP30664	Human sec	626	16	57.1	1371	8	ADP31646	Adp31646	Human sec
554	16	57.1	861	8	ADP31020	Human sec	627	16	57.1	1371	8	ADP30876	Adp30876	Human sec
555	16	57.1	861	8	ADP31021	Human sec	628	16	57.1	1380	8	ADP31566	Adp31566	Human sec
556	16	57.1	885	8	ADP31198	Human sec	629	16	57.1	1386	8	ADP31371	Adp31371	Human sec
557	16	57.1	887	8	ADP30554	Human sec	630	16	57.1	1387	8	ADP30946	Adp30946	Human sec
558	16	57.1	887	8	ADP30548	Human sec	631	16	57.1	1431	8	ADP31609	Adp31609	Human sec
559	16	57.1	888	8	ADP30971	Human sec	632	16	57.1	1440	8	ADG34533	Adg34533	Glucocort
560	16	57.1	890	8	ADP31059	Human sec	633	16	57.1	1456	8	ADP30923	Adp30923	Human sec
561	16	57.1	891	8	ADP31668	Human sec	634	16	57.1	1464	8	ADP31040	Adp31040	Human sec
562	16	57.1	897	8	ADP30914	Human sec	635	16	57.1	1476	7	ABO80681	Abob80681	Pseudomon
563	16	57.1	912	8	ADP31507	Human sec	636	16	57.1	1494	8	ADP31650	Adp31650	Human sec
564	16	57.1	918	8	ADP31459	Human sec	637	16	57.1	1506	8	ADP30586	Adp30586	Human sec
565	16	57.1	933	8	ADP31510	Human sec	638	16	57.1	1518	8	ADP31532	Adp31532	Human sec
566	16	57.1	936	8	ADP31486	Human sec	639	16	57.1	1566	8	ADP31407	Adp31407	Human sec
567	16	57.1	936	8	ADP31597	Human sec	640	16	57.1	1578	8	ADP30499	Adp30499	Human sec
568	16	57.1	939	8	ADP30726	Human sec	641	16	57.1	1584	8	ADP31405	Adp31405	Human sec
569	16	57.1	945	8	ADP31238	Human sec	642	16	57.1	1596	8	ADP31491	Adp31491	Human sec
570	16	57.1	945	8	ADP31237	Human sec	643	16	57.1	1602	9	ADZ66459	Adz66459	Human ARN
571	16	57.1	948	8	ADP30586	Human sec	644	16	57.1	1623	8	ADP30552	Adp30552	Human sec
572	16	57.1	966	8	ADP30745	Human sec	645	16	57.1	1623	8	ADP31653	Adp31653	Human sec
573	16	57.1	1001	8	ADP31365	Human sec	646	16	57.1	1629	8	ADP30945	Adp30945	Human sec
574	16	57.1	1002	8	ADG39639	Human pan	647	16	57.1	1647	8	ADP31052	Adp31052	Human sec
575	16	57.1	1002	8	ADP30866	Human sec	648	16	57.1	1670	8	ADP30927	Adp30927	Human sec
576	16	57.1	1008	8	ADP30721	Human sec	649	16	57.1	1719	8	ADP31137	Adp31137	Human sec
577	16	57.1	1010	8	ADP31296	Human sec	650	16	57.1	1746	8	ADP30992	Adp30992	Human sec
578	16	57.1	1017	6	AAO26719	SR protei	651	16	57.1	1749	8	ADP31408	Adp31408	Human sec
579	16	57.1	1017	8	ADP31268	Human sec	652	16	57.1	1771	8	ADP30674	Adp30674	Human sec
580	16	57.1	1023	8	ADP30994	Human sec	653	16	57.1	1782	8	ADP31391	Adp31391	Human sec
581	16	57.1	1030	8	ADP30913	Human sec	654	16	57.1	1789	8	ADP30962	Adp30962	Human sec
582	16	57.1	1050	8	ADP31370	Human sec	655	16	57.1	1794	8	ADP31176	Adp31176	Human sec
583	16	57.1	1056	8	ADP31082	Human sec	656	16	57.1	1815	8	ADP31601	Adp31601	Human sec
584	16	57.1	1057	8	ADP31592	Human sec	657	16	57.1	1827	8	ADP31170	Adp31170	Human sec
585	16	57.1	1062	8	ADP31369	Human sec	658	16	57.1	1848	8	ADP31372	Adp31372	Human sec
586	16	57.1	1065	8	ADP30661	Human sec	659	16	57.1	1875	8	ADP31664	Adp31664	Human sec
587	16	57.1	1065	8	ADP30973	Human sec	660	16	57.1	1878	8	ADP30740	Adp30740	Human sec
588	16	57.1	1065	8	ADP31482	Human sec	661	16	57.1	1917	6	ADA15719	Ada15719	C. elegan
589	16	57.1	1065	8	ADP31287	Human sec	662	16	57.1	1933	8	ADP30889	Adp30889	Human sec
590	16	57.1	1066	8	ADP30563	Human sec	663	16	57.1	1933	8	ADP30902	Adp30902	Human sec
591	16	57.1	1086	8	ADP31447	Human sec	664	16	57.1	1956	8	ADP31662	Adp31662	Human sec
592	16	57.1	1090	8	ADP30187	Human pol	665	16	57.1	1989	8	ADP31185	Adp31185	Human sec
593	16	57.1	1104	8	ADP31156	Human sec	666	16	57.1	1996	8	ADP31599	Adp31599	Human sec
594	16	57.1	1116	8	ADP31692	Human sec	667	16	57.1	2001	8	ADP31644	Adp31644	Human sec
595	16	57.1	1132	8	ADP30952	Human sec	668	16	57.1	2020	8	ADP31056	Adp31056	Human sec
596	16	57.1	1140	8	ADP31128	Human sec	669	16	57.1	2148	8	ADP30974	Adp30974	Human sec
597	16	57.1	1140	8	ADP31130	Human sec	670	16	57.1	2358	8	ADP31690	Adp31690	Human sec
598	16	57.1	1147	8	ADP30965	Human sec	671	16	57.1	2391	8	ADP31366	Adp31366	Human sec
599	16	57.1	1183	8	ADP30536	Human sec	672	16	57.1	2454	8	ADP30469	Adp30469	Human sec
600	16	57.1	1184	7	ADP30653	Human sec	673	16	57.1	2469	8	ADP31031	Adp31031	Human sec
601	16	57.1	1186	2	AAV05840	Banana x1	674	16	57.1	2514	8	ADP30727	Adp30727	Human sec
602	16	57.1	1194	8	ADP31406	Human sec	675	16	57.1	2535	8	ADP31146	Adp31146	Human sec
603	16	57.1	1199	8	ADP31044	Human sec	676	16	57.1	2547	8	ADP31665	Adp31665	Human sec
604	16	57.1	1200	9	ADP31603	Human sec	677	16	57.1	2611	8	ADP31027	Adp31027	Human sec
605	16	57.1	1200	9	ADZ76047	Mouse pre	678	16	57.1	2622	8	ADP31663	Adp31663	Human sec
606	16	57.1	1209	8	ADP30782	Human sec	679	16	57.1	2664	8	ADP31546	Adp31546	Human sec
607	16	57.1	1215	8	ADP30898	Human sec	680	16	57.1	2828	8	ADP30938	Adp30938	Human sec
608	16	57.1	1215	8	ADP30900	Human sec	681	16	57.1	2980	9	AEA36049	Aea36049	Maize Sta

682	16	57.1	3036	8	ADP31595	Adp31595	Human sec	755	15	53.6	111	7	ABO83484	AbO83484	Pseudomon
683	16	57.1	3046	8	ADQ10419	Adq10419	Human pol	756	15	53.6	112	5	ABE89458	AbB89458	Human pol
684	16	57.1	3144	8	ADP31544	Adp31544	Human sec	757	15	53.6	113	8	ADP31617	Adp31617	Human sec
685	16	57.1	3201	8	ADP31545	Adp31545	Human sec	758	15	53.6	118	4	AAU53541	Aau53541	Pseudomon
686	16	57.1	3465	8	ADP31234	Adp31234	Human sec	759	15	53.6	119	6	ABM50060	Abm50060	Pseudomon
687	16	57.1	3477	8	ADP30704	Adp30704	Human sec	760	15	53.6	123	7	ABO70463	AbO70463	Pseudomon
688	16	57.1	3579	8	ADP31098	Adp31098	Human sec	761	15	53.6	125	4	AAU47488	Aau47488	Pseudomon
689	16	57.1	3638	8	ADP30981	Adp30981	Human sec	762	15	53.6	126	6	ABM44007	Abm44007	Pseudomon
690	16	57.1	3907	5	ABG70822	Abg70822	Mouse myo	763	15	53.6	126	6	ABR57124	AbR57124	Camel ant
691	16	57.1	3907	6	ABG74190	Abg74190	Mouse myo	764	15	53.6	127	8	ADP31622	Adp31622	Human sec
692	16	57.1	4752	8	ADP30585	Adp30585	Human sec	765	15	53.6	130	7	ADP27051	Adp27051	Human adi
693	16	57.1	4752	8	ADP30651	Adp30651	Human sec	766	15	53.6	131	7	ABO66182	AbO66182	Klebsiell
694	16	57.1	5397	8	ADP31068	Adp31068	Human sec	767	15	53.6	134	7	ABO76972	AbO76972	Human sec
695	16	57.1	8973	8	ADP31119	Adp31119	Human sec	768	15	53.6	135	7	ABO83194	AbO83194	Pseudomon
696	15	53.6	12	6	ABG71386	Abg71386	Scaphyloc	769	15	53.6	135	7	ABO79475	AbO79475	Pseudomon
697	15	53.6	24	8	ADD95109	Add95109	PCR prime	770	15	53.6	135	8	ADP30703	Adp30703	Human sec
698	15	53.6	24	8	ADU59677	Adu59677	GHP-4 c10	771	15	53.6	136	7	ABO74222	AbO74222	Pseudomon
699	15	53.6	27	4	AAU13578	Aau13578	Peptide #	772	15	53.6	138	3	AAAB42027	AaB42027	Human ORF
700	15	53.6	27	4	ABR32506	Abb32506	Peptide #	773	15	53.6	139	7	ABO68581	AbO68581	Pseudomon
701	15	53.6	27	4	AAU25975	Aau25975	Peptide #	774	15	53.6	143	8	ADP31632	Adp31632	Human sec
702	15	53.6	27	4	ABR27361	Abb27361	Human pep	775	15	53.6	143	7	ABO77067	AbO77067	Pseudomon
703	15	53.6	27	4	ABR18013	Abb18013	Protein #	776	15	53.6	144	7	ABO75730	AbO75730	Pseudomon
704	15	53.6	27	4	AAU65718	Aau65718	Human bon	777	15	53.6	146	4	ABG02814	AbG02814	Novel hum
705	15	53.6	27	4	AAU53341	Aau53341	Human bra	778	15	53.6	150	8	AD145261	Ad145261	Rice isop
706	15	53.6	27	4	ABG47359	Abg47359	Human liv	779	15	53.6	152	8	ADP31605	Adp31605	Human sec
707	15	53.6	27	4	AAU01330	Aau01330	Peptide #	780	15	53.6	153	7	ABO83454	AbO83454	Pseudomon
708	15	53.6	27	5	ABG35348	Abg35348	Human pep	781	15	53.6	154	7	ABO75839	AbO75839	Pseudomon
709	15	53.6	34	4	AAU17951	Aau17951	Peptide #	782	15	53.6	156	7	ABO75667	AbO75667	Pseudomon
710	15	53.6	34	4	ABR36982	Abb36982	Peptide #	783	15	53.6	158	3	AAAG29118	AaG29118	Arabidops
711	15	53.6	41	8	ADP21826	Adp21826	Low densi	784	15	53.6	163	7	ADD78226	AdD78226	Human CGD
712	15	53.6	61	4	AAU51383	Aau51383	Proteinib	785	15	53.6	163	7	ABO80383	AbO80383	Pseudomon
713	15	53.6	61	6	ABM47992	Abm47992	Proteinib	786	15	53.6	165	8	ADP30587	Adp30587	Human sec
714	15	53.6	69	6	AAU62232	Aau62232	Proteinib	787	15	53.6	165	8	ADP31041	Adp31041	Human sec
715	15	53.6	69	6	ABM58751	Abm58751	Proteinib	788	15	53.6	168	8	ADP30841	Adp30841	Human sec
716	15	53.6	70	6	ABR04032	Abbr04032	Human mus	789	15	53.6	169	7	ABO74424	AbO74424	Pseudomon
717	15	53.6	70	6	ABU13336	Abu13336	Novel hum	790	15	53.6	171	8	ABO70973	AbO70973	Pseudomon
718	15	53.6	72	7	ABO66669	AbO66669	Klebsiell	791	15	53.6	173	8	ADY10989	AdY10989	Plant ful
719	15	53.6	72	7	ABR37637	Abb37637	Peptide #	792	15	53.6	174	8	ADP31530	Adp31530	Human sec
720	15	53.6	76	4	AAU70732	Aau70732	Human bon	793	15	53.6	177	7	ABO70036	AbO70036	Pseudomon
721	15	53.6	76	4	ABG52439	Abg52439	Human liv	795	15	53.6	177	8	ADP30755	Adp30755	Human sec
722	15	53.6	76	4	AAU06152	Aau06152	Peptide #	796	15	53.6	179	7	ABO78051	AbO78051	Pseudomon
723	15	53.6	77	4	AAU91132	Aau91132	Human imm	797	15	53.6	183	4	AAU54675	Aau54675	Proteinib
724	15	53.6	80	4	AAU59349	Aau59349	Proteinib	798	15	53.6	183	6	ABM51194	AbM51194	Pseudomon
725	15	53.6	80	6	ABM55868	Abm55868	Proteinib	799	15	53.6	184	7	ABO77341	AbO77341	Pseudomon
726	15	53.6	80	7	ABO71261	AbO71261	Pseudomon	800	15	53.6	187	9	ABE03464	Abe03464	Mycobacte
727	15	53.6	83	4	AAU53165	Aau53165	Proteinib	801	15	53.6	187	9	ABE79367	Abe79367	Novel M.
728	15	53.6	83	6	ABM49684	Abm49684	Proteinib	802	15	53.6	188	6	ABM69571	AbM69571	Phototrab
729	15	53.6	85	4	AAU46688	Aau46688	Proteinib	803	15	53.6	188	7	ABO81563	AbO81563	Pseudomon
730	15	53.6	85	6	ABM43387	Abm43387	Proteinib	804	15	53.6	189	8	ADP31641	Adp31641	Human sec
731	15	53.6	86	7	ABO78259	AbO78259	Pseudomon	805	15	53.6	189	8	ADP30641	Adp30641	Human sec
732	15	53.6	86	7	AAU41562	Aau41562	Proteinib	806	15	53.6	192	7	AD163118	Ad163118	Human epo
733	15	53.6	87	4	ABG07613	Abg07613	Novel hum	807	15	53.6	192	8	ADP30573	Adp30573	Human sec
734	15	53.6	87	6	ABM38081	Abm38081	Proteinib	808	15	53.6	192	8	ADP30575	Adp30575	Human sec
735	15	53.6	88	7	ABO76533	AbO76533	Pseudomon	809	15	53.6	192	8	ADX73861	AdX73861	Plant ful
736	15	53.6	89	4	ABG27446	Abg27446	Novel hum	810	15	53.6	194	7	ABO78166	AbO78166	Pseudomon
737	15	53.6	89	6	ADK34238	Adk34238	Actinectoda	811	15	53.6	194	8	ADP31626	Adp31626	Human sec
738	15	53.6	90	6	AAU50672	Aau50672	Proteinib	812	15	53.6	195	8	ADP30696	Adp30696	Human sec
739	15	53.6	95	4	AAU40381	Aau40381	Proteinib	813	15	53.6	198	8	ADP30493	Adp30493	Human sec
740	15	53.6	95	4	ABM47191	Abm47191	Proteinib	814	15	53.6	198	8	ADP30492	Adp30492	Human sec
741	15	53.6	95	6	ABM36900	Abm36900	Proteinib	815	15	53.6	198	8	ADP30477	Adp30477	Human sec
742	15	53.6	101	4	AAU83737	Aau83737	Human imm	816	15	53.6	198	8	ADP30481	Adp30481	Human sec
743	15	53.6	102	4	AAU61573	Aau61573	Proteinib	817	15	53.6	200	7	ABO73357	AbO73357	Human sec
744	15	53.6	102	6	ABM58092	Abm58092	Proteinib	818	15	53.6	200	7	ADP30686	Adp30686	Human sec
745	15	53.6	103	4	AAU48984	Aau48984	Proteinib	819	15	53.6	201	3	AAAG29117	AaG29117	Arabidops
746	15	53.6	103	6	ABM45503	Abm45503	Proteinib	820	15	53.6	204	8	ADP30488	Adp30488	Human sec
747	15	53.6	103	7	ABO84304	AbO84304	Pseudomon	821	15	53.6	204	8	ADP31421	Adp31421	Human sec
748	15	53.6	106	2	AAU29187	Aau29187	Amino aci	822	15	53.6	204	8	ADP31422	Adp31422	Human sec
749	15	53.6	107	4	AAU43314	Aau43314	Proteinib	823	15	53.6	204	8	ADP30489	Adp30489	Human sec
750	15	53.6	107	5	ABP64480	Abp64480	Human ORF	824	15	53.6	204	8	ADP31424	Adp31424	Human sec
751	15	53.6	108	3	AAAB21190	AaB21190	Exol4 par	825	15	53.6	204	8	ADP31420	Adp31420	Human sec
752	15	53.6	110	9	ABM95190	Abm95190	M. xanthu	827	15	53.6	204	8	ADP30490	Adp30490	Human sec

828	15	53.6	207	7	ABO71507	AbO71507 Pseudomon	901	15	53.6	339	8	ADP30519	Adp30519 Human sec
829	15	53.6	208	7	ABO75819	AbO75819 Pseudomon	902	15	53.6	339	8	ADP30702	Adp30702 Human sec
830	15	53.6	208	7	ABO78684	AbO78684 Pseudomon	903	15	53.6	348	8	ADP31374	Adp31374 Human sec
831	15	53.6	210	8	ADP30986	Adp30986 Human sec	904	15	53.6	349	5	ABO86605	AbO86605 Hlrudin p
832	15	53.6	210	8	ADP30960	Adp30960 Human sec	905	15	53.6	350	4	ABO86635	AbO86635 Drosophil
833	15	53.6	212	2	AAW82678	X. cumber	906	15	53.6	357	8	ADP30961	Adp30961 Human sec
834	15	53.6	212	2	AAW82685	L. cornic	907	15	53.6	361	3	AAO42213	AAO42213 Human ORF
835	15	53.6	212	2	AAW82677	X. cumber	908	15	53.6	362	8	ADN99707	ADN99707 Novel hum
836	15	53.6	213	2	AAW82677	X. cumber	909	15	53.6	362	8	ADN99707	ADN99707 Novel hum
837	15	53.6	216	7	ABO84141	AbO84141 C glutami	910	15	53.6	362	8	ABM61662	ABM61662 Human PRO
838	15	53.6	217	7	ABO79989	AbO79989 Pseudomon	911	15	53.6	366	8	ADP31670	Adp31670 Human sec
839	15	53.6	218	7	ABO79989	AbO79989 Pseudomon	912	15	53.6	366	8	ADP30582	Adp30582 Human sec
840	15	53.6	222	4	ABG27806	Novel hum	913	15	53.6	377	8	ADP31669	Adp31669 Human sec
841	15	53.6	222	7	ABO77741	AbO77741 Pseudomon	914	15	53.6	382	8	ADP31216	Adp31216 Human sec
842	15	53.6	222	8	ADP31108	Adp31108 Human sec	915	15	53.6	387	4	ABO71334	ABO71334 Drosophil
843	15	53.6	223	7	ABO77463	AbO77463 Pseudomon	916	15	53.6	387	8	ADP31158	Adp31158 Human sec
844	15	53.6	224	8	ADP31480	Adp31480 Human sec	917	15	53.6	390	8	ADP31584	Adp31584 Human sec
845	15	53.6	225	8	ADP31480	Adp31480 Human sec	918	15	53.6	399	8	ADP31367	Adp31367 Human sec
846	15	53.6	226	4	ABM52385	ABM52385 Human CD8	919	15	53.6	399	8	ADP31317	Adp31317 Human sec
847	15	53.6	227	4	ABM52385	ABM52385 Human CD8	920	15	53.6	399	8	ADP31313	Adp31313 Human sec
848	15	53.6	228	8	ADP31334	Adp31334 Human sec	921	15	53.6	414	8	ADP30979	Adp30979 Human sec
849	15	53.6	234	8	ADP31480	Adp31480 Human sec	922	15	53.6	417	8	ADP31432	Adp31432 Human sec
850	15	53.6	234	8	ADP31481	Adp31481 Human sec	923	15	53.6	420	8	ADN20155	ADN20155 Bacterial
851	15	53.6	237	8	ADP30792	Adp30792 Human sec	924	15	53.6	423	8	ADP30819	Adp30819 Human sec
852	15	53.6	239	7	ABO75548	AbO75548 Pseudomon	925	15	53.6	423	8	ADP31323	Adp31323 Human sec
853	15	53.6	240	6	ABR47527	Breast ca	926	15	53.6	432	8	ADP31365	Adp31365 Human sec
854	15	53.6	240	6	ABR47527	Breast ca	927	15	53.6	435	7	ABO68739	ABO68739 Pseudomon
855	15	53.6	240	7	ADN38776	Cancer/an	928	15	53.6	442	8	ADP31134	Adp31134 Human sec
856	15	53.6	240	7	ABO78355	Pseudomon	929	15	53.6	443	3	AAO13372	AAO13372 Arabidops
857	15	53.6	243	8	ADP30983	Adp30983 Human sec	930	15	53.6	456	8	ADP31224	Adp31224 Human sec
858	15	53.6	247	7	ABO72974	AbO72974 Pseudomon	931	15	53.6	463	7	ABO71740	ABO71740 Pseudomon
859	15	53.6	248	7	ABO71415	AbO71415 Pseudomon	932	15	53.6	471	8	ADP31567	Adp31567 Human sec
860	15	53.6	248	8	ADP30662	Adp30662 Human sec	933	15	53.6	471	8	ADP30870	Adp30870 Human sec
861	15	53.6	249	8	ADP30754	Adp30754 Human sec	934	15	53.6	473	8	ADP30894	Adp30894 Human sec
862	15	53.6	249	8	ADP31315	Adp31315 Human sec	935	15	53.6	473	8	ADP30728	Adp30728 Human sec
863	15	53.6	250	7	ABO83401	AbO83401 Pseudomon	936	15	53.6	474	8	ADP31157	Adp31157 Human sec
864	15	53.6	251	7	ABO73872	Pseudomon	937	15	53.6	478	8	ADP31007	Adp31007 Human sec
865	15	53.6	254	8	ADP31396	Adp31396 Human sec	938	15	53.6	484	8	ADP31628	Adp31628 Human sec
866	15	53.6	268	3	AAO29116	AAO29116 Arabidops	939	15	53.6	485	3	AAO59032	AAO59032 Breast an
867	15	53.6	268	7	ADP30662	Adp30662 Ptalectes	940	15	53.6	491	4	ABO10165	ABO10165 Novel hum
868	15	53.6	268	8	ADP31697	Adp31697 Human sec	941	15	53.6	496	4	ABO66085	ABO66085 Drosophil
869	15	53.6	268	8	ADP31697	Adp31697 Human sec	942	15	53.6	502	2	AAO2762	AAO2762 Thyroid N
870	15	53.6	268	8	ADP31321	Adp31321 Plant ful	943	15	53.6	504	8	ADP31624	Adp31624 Human sec
871	15	53.6	270	8	ADP31321	Adp31321 Human sec	944	15	53.6	513	8	ADP31681	Adp31681 Human sec
872	15	53.6	270	7	ABO71564	AbO71564 Human sec	945	15	53.6	516	8	ADP30871	Adp30871 Human sec
873	15	53.6	271	7	ABO77489	AbO77489 Pseudomon	946	15	53.6	520	7	ADP31845	Adp31845 Human nov
874	15	53.6	274	8	ADP31679	Adp31679 Human sec	947	15	53.6	520	7	ADP305027	ADP305027 Human pto
875	15	53.6	277	8	ADP30503	Adp30503 Human sec	948	15	53.6	522	8	ADP31081	Adp31081 Human sec
876	15	53.6	279	8	ADP31509	Adp31509 Human sec	949	15	53.6	522	8	ADP31018	Adp31018 Human sec
877	15	53.6	279	8	ADP30846	Adp30846 Human sec	950	15	53.6	524	8	ADP31018	Adp31018 Human sec
878	15	53.6	279	8	ADP30846	Adp30846 Human sec	951	15	53.6	524	4	AAU07370	AAU07370 G protein
879	15	53.6	287	2	AAO73013	AAO73013 Punctiona	952	15	53.6	525	5	ABO60120	ABO60120 Human DIT
880	15	53.6	288	2	AAO73013	AAO73013 Punctiona	953	15	53.6	526	6	ADP31054	Adp31054 Human sec
881	15	53.6	289	2	AAO73013	AAO73013 Punctiona	954	15	53.6	527	6	ABO43310	ABO43310 Human lrp
882	15	53.6	293	3	AAO73013	AAO73013 Punctiona	955	15	53.6	534	7	ABO83340	ABO83340 Pseudomon
883	15	53.6	294	4	ABO83340	ABO83340 Pseudomon	956	15	53.6	541	6	ABO25731	ABO25731 Aspergill
884	15	53.6	294	4	ABO83340	ABO83340 Pseudomon	957	15	53.6	543	8	ADP31506	Adp31506 Human sec
885	15	53.6	300	8	ADP30616	Adp30616 Human sec	958	15	53.6	543	8	ADP30872	Adp30872 Human sec
886	15	53.6	300	8	ADP30616	Adp30616 Human sec	959	15	53.6	543	8	ADP31393	Adp31393 Human sec
887	15	53.6	306	8	ADP31205	Adp31205 Human sec	960	15	53.6	552	8	ADP30875	Adp30875 Human sec
888	15	53.6	311	3	AAO34889	AAO34889 Human sec	961	15	53.6	555	8	ADP31416	Adp31416 Human sec
889	15	53.6	311	4	AAO34889	AAO34889 Human sec	962	15	53.6	555	8	ADP31417	Adp31417 Human sec
890	15	53.6	312	4	AAO34889	AAO34889 Human sec	963	15	53.6	558	8	ADP31257	Adp31257 Human sec
891	15	53.6	312	8	ADP31454	Adp31454 Human sec	964	15	53.6	564	6	ABO26331	ABO26331 Aspergill
892	15	53.6	318	8	ADP31066	Adp31066 Human sec	965	15	53.6	564	8	ADP31194	Adp31194 Human sec
893	15	53.6	319	7	ABO72910	AbO72910 Pseudomon	966	15	53.6	567	8	ADP31395	Adp31395 Human sec
894	15	53.6	321	3	AAO13373	AAO13373 Arabidops	967	15	53.6	569	7	ADP31395	Adp31395 Human sec
895	15	53.6	326	8	ADP31254	Adp31254 Arabidops	968	15	53.6	574	8	ADP31372	Adp31372 Rice ABC
896	15	53.6	332	7	ABO83940	ABO83940 Human GPC	969	15	53.6	575	8	ADP31143	Adp31143 Human sec
897	15	53.6	337	8	ADP31387	Adp31387 Human sec	970	15	53.6	583	8	ADP30553	Adp30553 Human sec
898	15	53.6	338	8	ABO58310	ABO58310 Human gen	971	15	53.6	585	8	ADP31392	Adp31392 Human sec
900	15	53.6	339	8	ADP30892	ADP30892 Human sec	972	15	53.6	585	8	ADP31445	Adp31445 Human sec
							973	15	53.6	585	8	ADP31445	Adp31445 Human sec

974	15	53.6	587	4	AB867608	Abb57608 Drosophil
975	15	53.6	588	8	ADP30877	Adp30877 Human sec
976	15	53.6	588	8	ADP31680	Adp31680 Human sec
977	15	53.6	588	8	ADP31699	Adp31699 Human sec
978	15	53.6	591	2	AAY15228	Aay15228 Human rec
979	15	53.6	594	8	ADP30895	Adp30895 Human sec
980	15	53.6	603	8	ADP31150	Adp31150 Human sec
981	15	53.6	603	8	ADP31645	Adp31645 Human sec
982	15	53.6	605	8	ADP30507	Adp30507 Human sec
983	15	53.6	608	8	ADP04735	Adp04735 Sea squit
984	15	53.6	609	8	ADP31263	Adp31263 Human sec
985	15	53.6	612	7	AB069512	Ab069512 Pseudomon
986	15	53.6	612	8	ADP31064	Adp31064 Human sec
987	15	53.6	615	8	ADP31361	Adp31361 Human sec
988	15	53.6	615	8	ADP30803	Adp30803 Human sec
989	15	53.6	615	8	ADP31360	Adp31360 Human sec
990	15	53.6	617	8	ADP31657	Adp31657 Human sec
991	15	53.6	621	8	ADP30896	Adp30896 Human sec
992	15	53.6	621	8	ADP31147	Adp31147 Human sec
993	15	53.6	632	7	AB076798	Ab076798 Pseudomon
994	15	53.6	637	8	ADP31397	Adp31397 Human sec
995	15	53.6	638	8	ADP30513	Adp30513 Human sec
996	15	53.6	640	8	ADP30977	Adp30977 Human sec
997	15	53.6	642	8	ADR86228	Adr86228 Aspergill
998	15	53.6	645	8	ADP31067	Adp31067 Human sec
999	15	53.6	648	6	ABR43306	Adr43306 Human l1p
1000	15	53.6	651	8	ADP31543	Adp31543 Human sec

ALIGNMENTS

```

RESULT 1
AB082562
ID      AB082562 standard; protein; 104 AA.
AC      ABO82562;
XX      29-JUL-2004 (first entry)
DT      XX
DE      Pseudomonas aeruginosa polypeptide #14737.
XX      XX
KW      Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
OS      Pseudomonas aeruginosa.
XX      XX
PN      US6551795-B1.
XX      XX
PD      22-APR-2003.
XX      XX
PF      18-FEB-1999; 99US-00252991.
XX      XX
PR      18-FEB-1998; 98US-0074788P.
PR      27-JUL-1998; 98US-0094190P.
XX      XX
PA      (GENO-) GENOME THERAPEUTICS CORP.
XX      XX
PI      Rubenfield MJ, Nolling J, DeLoughery C, Bush D;
DR      WPI: 2003-615309/58.
XX      N-PSDB: ABD16133.
PT      Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT      useful as molecular targets for diagnostics, prophylaxis and treatment of
PT      pathological conditions resulting from bacterial infection.
XX      XX
PS      Disclosure; SEQ ID NO 31308; 455pp; English.
XX      XX
CC      The invention relates to Pseudomonas aeruginosa polypeptides and the
CC      polynucleotides encoding them. The sequences are useful in diagnosis and
CC      therapy of pathological conditions, as molecular targets for diagnostics,
CC      prophylaxis and treatment of pathological conditions resulting from a
CC      bacterial infection, for evaluating a compound, such as a polypeptide,
```

```

CC      for the ability to bind a P. aeruginosa nucleic acid, as components of
CC      effective antibacterial targets, as targets for antibacterial drugs,
CC      including anti-P. aeruginosa drugs, as templates for recombinant
CC      production of P. aeruginosa-derived peptides or polypeptides, as target
CC      components for diagnosis and/or treatment of P. aeruginosa-caused
CC      infection, and in detection of P. aeruginosa sequences or other sequences
CC      of Pseudomonas species using biochip technology. Sequences AB067826-
CC      AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
CC      sequence data for this patent did not form part of the printed
CC      specification but was obtained in electronic format from USPTO at
CC      segdata.uspto.gov/sequence.html
XX      XX
SQ      Sequence 104 AA;
XX      XX
Query Match      64.3%; Score 18; DB 7; Length 104;
Best Local Similarity 16.7%; Pred. No. 4.9e-05;
Matches      2; Conservative      0; Mismatches      10; Indels      0; Gaps      0;
QY      1 CXXXXXXXXXXC 12
DB      56 CSATASASTTC 67

RESULT 2
ADP30723
ID      ADP30723 standard; protein; 110 AA.
AC      ADP30723;
XX      XX
DT      12-AUG-2004 (first entry)
XX      XX
DE      Human secreted protein SEQ ID #1490.
XX      XX
KW      Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW      cancer; inflammatory; immune; human sedreted protein.
XX      XX
OS      Homo sapiens.
XX      XX
PN      WO2004035732-A2.
XX      XX
PD      29-APR-2004.
XX      XX
PF      28-AUG-2003; 2003WO-US026780.
XX      XX
PR      29-AUG-2002; 2002US-0406576P.
PR      29-AUG-2002; 2002US-0406579P.
PR      29-AUG-2002; 2002US-0406585P.
PR      29-AUG-2002; 2002US-0406588P.
PR      29-AUG-2002; 2002US-0406608P.
PR      29-AUG-2002; 2002US-0406611P.
PR      29-AUG-2002; 2002US-0406612P.
PR      29-AUG-2002; 2002US-0406616P.
PR      29-AUG-2002; 2002US-0406640P.
PR      29-AUG-2002; 2002US-0406642P.
PR      29-AUG-2002; 2002US-0406646P.
PR      29-AUG-2002; 2002US-0406653P.
PR      29-AUG-2002; 2002US-0406655P.
PR      29-AUG-2002; 2002US-0406666P.
PR      17-SEP-2002; 2002US-0410946P.
PR      17-SEP-2002; 2002US-0410947P.
PR      17-SEP-2002; 2002US-0410948P.
PR      17-SEP-2002; 2002US-0410949P.
PR      17-SEP-2002; 2002US-0410951P.
PR      17-SEP-2002; 2002US-0410957P.
PR      17-SEP-2002; 2002US-0410957P.
PR      17-SEP-2002; 2002US-0410958P.
PR      17-SEP-2002; 2002US-0410959P.
PR      17-SEP-2002; 2002US-0410960P.
PR      17-SEP-2002; 2002US-0410961P.
PR      17-SEP-2002; 2002US-0410962P.
PR      17-SEP-2002; 2002US-0411019P.
PR      17-SEP-2002; 2002US-0411022P.
PR      17-SEP-2002; 2002US-0411023P.
PR      17-SEP-2002; 2002US-0411024P.
```



PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
PR (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX MPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 2721; 428bp; English.  
XX  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
XX  
SQ Sequence 110 AA;

Query Match 64.3%; Score 18; DB 8; Length 110;  
Best Local Similarity 16.7%; Pred. No. 5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 93 CATMAATATAAC 104

RESULT 3  
ADP31638  
ID ADP31638 standard; protein; 135 AA.  
XX  
XX ADP31638;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2405.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX Cancer; inflammatory; immune; human secreted protein.  
XX Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
XX  
XX  
XX 28-AUG-2003; 2003MO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406555P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411101P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 18-APR-2003; 2003US-0463700P.  
XX 18-APR-2003; 2003US-0463708P.  
XX 18-APR-2003; 2003US-0463716P.  
XX 18-APR-2003; 2003US-0463732P.  
XX 02-MAY-2003; 2003US-0467199P.  
XX 02-MAY-2003; 2003US-0467201P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 19-MAY-2003; 2003US-0471306P.



PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
PS Disclosure; SEQ ID NO 22030; 455bp; English.  
XX  
XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-*P. aeruginosa* drugs, as templates for recombinant  
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-  
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html)  
XX  
XX Sequence 178 AA;  
SQ  
Query Match 64.3%; Score 18; DB 7; Length 178;  
Best Local Similarity 16.7%; Pred. No. 5.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXXC 12  
Db 92 CATATATASSTSC 103  
RESULT 6  
ADP31527  
ID ADP31527 standard; protein; 264 AA.  
XX  
XX ADP31527;  
AC  
XX  
DT 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #2294.  
DE  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX  
XX WO2004035732-A2.  
XX  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485235P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kothakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1, SEQ ID NO 3525; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.

```
XX SQ Sequence 264 AA;
Query Match 64.3%; Score 18; DB 8; Length 264;
Best Local Similarity 16.7%; Pred. No. 6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 109 CAAATATTTC 120

RESULT 7
ADP31435
ID ADP31435 standard; protein; 270 AA.
XX
XX ADP31435;
AC 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #2202.
DE
XX
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
OS Homo sapiens.
XX
XX WO2004035732-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 28-AUG-2003; 2003WO-US026780.
PF
XX
XX 29-AUG-2002; 2002US-0406576P.
PR
XX 29-AUG-2002; 2002US-0406579P.
PR
XX 29-AUG-2002; 2002US-0406585P.
PR
XX 29-AUG-2002; 2002US-0406588P.
PR
XX 29-AUG-2002; 2002US-0406608P.
PR
XX 29-AUG-2002; 2002US-0406611P.
PR
XX 29-AUG-2002; 2002US-0406612P.
PR
XX 29-AUG-2002; 2002US-0406616P.
PR
XX 29-AUG-2002; 2002US-0406640P.
PR
XX 29-AUG-2002; 2002US-0406642P.
PR
XX 29-AUG-2002; 2002US-0406646P.
PR
XX 29-AUG-2002; 2002US-0406653P.
PR
XX 29-AUG-2002; 2002US-0406655P.
PR
XX 29-AUG-2002; 2002US-0406666P.
PR
XX 17-SEP-2002; 2002US-0410946P.
PR
XX 17-SEP-2002; 2002US-0410947P.
PR
XX 17-SEP-2002; 2002US-0410948P.
PR
XX 17-SEP-2002; 2002US-0410949P.
PR
XX 17-SEP-2002; 2002US-0410953P.
PR
XX 17-SEP-2002; 2002US-0410957P.
PR
XX 17-SEP-2002; 2002US-0410958P.
PR
XX 17-SEP-2002; 2002US-0410959P.
PR
XX 17-SEP-2002; 2002US-0410960P.
PR
XX 17-SEP-2002; 2002US-0410961P.
PR
XX 17-SEP-2002; 2002US-0410962P.
PR
XX 17-SEP-2002; 2002US-0411019P.
PR
XX 17-SEP-2002; 2002US-0411022P.
PR
XX 17-SEP-2002; 2002US-0411023P.
PR
XX 17-SEP-2002; 2002US-0411024P.
PR
XX 17-SEP-2002; 2002US-0411032P.
PR
XX 17-SEP-2002; 2002US-0411035P.
PR
XX 17-SEP-2002; 2002US-0411037P.
PR
XX 17-SEP-2002; 2002US-0411041P.
PR
XX 17-SEP-2002; 2002US-0411045P.
PR
XX 17-SEP-2002; 2002US-0411046P.
PR
XX 17-SEP-2002; 2002US-0411048P.
PR
XX 17-SEP-2002; 2002US-0411052P.
PR
XX 17-SEP-2002; 2002US-0411055P.
PR
XX 17-SEP-2002; 2002US-0411073P.
PR
XX 17-SEP-2002; 2002US-0411082P.
```

```
PR 17-SEP-2002; 2002US-0411101P.
PR 17-SEP-2002; 2002US-0411111P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493572P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kohakota S, Haishan L, Linnemann T;
PI Pierce K, Wang Y, Wong UGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX PS Claim 1, SEQ ID NO 3433; 428bp; English.
XX
XX CC The present invention relates to an isolated nucleic acid molecule
XX encoding a polypeptide which is believed to be cytostatic,
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX composition and methods are useful for diagnosing, preventing and
XX treating diseases such as proliferative (e.g. cancer), inflammatory,
XX immune, metabolic, genetic, bacterial and viral diseases. The present
XX sequence represents a human secreted protein. The present sequence is
XX available on WIPOMEB and is not in the specification.
XX
XX SQ Sequence 270 AA;
Query Match 64.3%; Score 18; DB 8; Length 270;
Best Local Similarity 16.7%; Pred. No. 6e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
Db 179 CTATATTTAAC 150

RESULT 8
ABB65879
ID ABB65879 standard; protein; 307 AA.
XX
XX ABB65879;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 24429.
DE
XX
```

KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO20017042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PP 23-MAR-2001; 2001WO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX N-PSDB; ABL09982.  
DR  
XX WPI; 2001-656860/75.  
XX  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT gene from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
XX  
PS Disclosure; SEQ ID NO 24429; 21np + Sequence listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB1616175-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (AB573737-  
CC AB872072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 307 AA;  
XX  
Query Match 64.3%; Score 18; DB 4; Length 307;  
Best Local Similarity 16.7%; Pred. No. 6.2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXXC 12  
DB 44 CTTTTTTTTTTC 55  
RESULT 9  
ADP31223  
ID ADP31223 standard; protein; 357 AA.  
XX  
XX ADP31223;  
AC  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1990.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
XX  
XX 29-APR-2004.  
PD  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411025P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476009P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486896P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX  
PI William LF, Chu K, Lee B, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Hsieh L, Linemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
DR WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3221; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.  
XX  
SQ Sequence 357 AA;  
Query Match 64.3%; Score 18; DB 8; Length 357;  
Best Local Similarity 16.7%; Pred. No. 6.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
CY 1 CXXXXXXXXXXC 12  
DB 175 CTTAATTTC 186  
RESULT 10  
ADP31132  
ID ADP31132 standard; protein; 615 AA.  
XX  
AC ADP31132;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1899.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PM WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 09-JUN-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 08-JUL-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 14-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 15-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493373P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RF, Huang MM, Kotlakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3130; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.  
XX  
SQ Sequence 615 AA;  
Query Match 64.3%; Score 18; DB 8; Length 615;  
Best Local Similarity 16.7%; Pred. No. 7.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
DB 175 CATTATTTTAAAC 186

## RESULT 11

AAU51580  
ID AAU51580 standard; protein; 627 AA.

AAU51580;

27-FEB-2002 (first entry)

Propionibacterium acnes immunogenic protein #12476.

SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; ELISA, inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteopathic; neuroprotectant.

Propionibacterium acnes.

WO200181581-A2.

01-NOV-2001.

20-APR-2001; 2001WO-US012865.

21-APR-2000; 2000US-0199047P.

02-JUN-2000; 2000US-0208841P.

07-JUL-2000; 2000US-0216747P.

(CORI-) CORIXA CORP.

Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhactia A;

L'maisonneuve J, Zhang Y, Jen S, Carter D;

WPI; 2001-616774/71.

N-PSDB; AAS59551.

Example 1; SEQ ID NO 12775; 1069pp; English.

Sequences AAU51580-17 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in

the treatment, prevention and diagnosis of medical conditions caused by

P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.

P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory

lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a

sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The

polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to

downregulate expression and activity of P. acnes polypeptides and

therefore treat P. acnes infections. The antibodies may also be used as

diagnostic agents for determining P. acnes presence, for example, by

enzyme linked immunosorbent assay (ELISA). Note: The sequence data for this patent did not form part of the printed specification, but was

obtained in electronic format directly from WIPO at

ftp.wipo.int/pub/published\_pct\_sequences

Sequence 627 AA; Query Match 64.3%; Score 18; DB 4; Length 627; Best Local Similarity 16.7%; Pred. No. 7.2e-05; Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
DB 46 CATTTSATNSC 57

## RESULT 12

ABM48099  
ID ABM48099 standard; protein; 627 AA.

ABM48099;

20-OCT-2003 (first entry)

Propionibacterium acnes predicted ORF-encoded polypeptide #12775.

Acne vulgaris; antisborrheic; dermatological; antibacterial; immunostimulant; immune response; vaccine.

Propionibacterium acnes.

WO2003033515-A1.

24-APR-2003.

11-OCT-2002; 2002WO-US032727.

15-OCT-2001; 2001US-00978825.

(CORI-) CORIXA CORP.

Mitcham JL, Skeiky YAM, Persing DH, Bhactia A, Maisonneuve JL;

Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

Barth B, Vallieve-Douglas J;

WPI; 2003-381789/36.

N-PSDB; ACF64480.

New Propionibacterium acnes polypeptides and polynucleotides encoding the

polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

or for stimulating an immune response specific for a P. acnes protein.

Example 1; SEQ ID NO 12775; 1481pp; English.

The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

encoding a Propionibacterium acnes protein. The invention also relates to

polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

immunogenic fragments of P. acnes polypeptides. The invention

additionally encompasses expression vectors and host cells comprising a

polynucleotide of the invention; antibodies against polypeptides of the

invention; fusion proteins comprising a polypeptide of the invention; a

method for stimulating an immune response specific for a P. acnes

polypeptide and an isolated T cell population comprising T cells prepared

via this method; a vaccine composition (comprising P. acnes polypeptides,

polynucleotides, antibodies, fusion proteins, T cell populations, or

antigen-presenting cells that express the polypeptide); a method and kit

for detecting or determining the presence or absence of P. acnes in a

patient; and a method for inhibiting the development of P. acnes in a

patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

proteins, T cell populations or antigen-presenting cells that express the

polypeptides are useful for diagnosing, preventing or treating acne

vulgaris, or for stimulating an immune response specific for a P. acnes

protein. The polynucleotides can also be used as probes or primers for

nucleic acid hybridization. The vaccine composition is useful for the

stimulation of an immune response against P. acnes, or for treating acne,

and the kit is useful for performing a diagnostic assay. The present

sequence represents a polypeptide predicted to be encoded by an ORF (open

reading frame) contained within the P. acnes polynucleotides of the

invention. Note: The sequence data for this patent did not form part of

the printed specification, but was obtained in electronic format directly

from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 627 AA;

```
Query Match          64.3%; Score 18; DB 6; Length 627;
Best Local Similarity 16.7%; Pred. No. 7.2e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY      1 CXXXXXXXXXXC 12
Db      46 CAATTTTATSC 57

RESULT 13
ADP31142
ID ADP31142 standard; protein; 669 AA.
XX
AC ADP31142;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1909.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
XX
OS Homo sapiens.
XX
PN W02004035732-A2.
XX
PD 29-APR-2004.
XX
PF 28-AUG-2003; 2003WO-US026780.
XX
PR 29-AUG-2002; 2002US-0406576P.
XX
PR 29-AUG-2002; 2002US-0406579P.
XX
PR 29-AUG-2002; 2002US-0406585P.
XX
PR 29-AUG-2002; 2002US-0406588P.
XX
PR 29-AUG-2002; 2002US-0406608P.
XX
PR 29-AUG-2002; 2002US-0406611P.
XX
PR 29-AUG-2002; 2002US-0406612P.
XX
PR 29-AUG-2002; 2002US-0406616P.
XX
PR 29-AUG-2002; 2002US-0406640P.
XX
PR 29-AUG-2002; 2002US-0406642P.
XX
PR 29-AUG-2002; 2002US-0406646P.
XX
PR 29-AUG-2002; 2002US-0406653P.
XX
PR 29-AUG-2002; 2002US-0406655P.
XX
PR 29-AUG-2002; 2002US-0406666P.
XX
PR 17-SEP-2002; 2002US-0410946P.
XX
PR 17-SEP-2002; 2002US-0410947P.
XX
PR 17-SEP-2002; 2002US-0410948P.
XX
PR 17-SEP-2002; 2002US-0410949P.
XX
PR 17-SEP-2002; 2002US-0410952P.
XX
PR 17-SEP-2002; 2002US-0410953P.
XX
PR 17-SEP-2002; 2002US-0410957P.
XX
PR 17-SEP-2002; 2002US-0410958P.
XX
PR 17-SEP-2002; 2002US-0410959P.
XX
PR 17-SEP-2002; 2002US-0410960P.
XX
PR 17-SEP-2002; 2002US-0410961P.
XX
PR 17-SEP-2002; 2002US-0410962P.
XX
PR 17-SEP-2002; 2002US-0411019P.
XX
PR 17-SEP-2002; 2002US-0411022P.
XX
PR 17-SEP-2002; 2002US-0411023P.
XX
PR 17-SEP-2002; 2002US-0411024P.
XX
PR 17-SEP-2002; 2002US-0411032P.
XX
PR 17-SEP-2002; 2002US-0411035P.
XX
PR 17-SEP-2002; 2002US-0411037P.
XX
PR 17-SEP-2002; 2002US-0411041P.
XX
PR 17-SEP-2002; 2002US-0411045P.
XX
PR 17-SEP-2002; 2002US-0411046P.
XX
PR 17-SEP-2002; 2002US-0411048P.
XX
PR 17-SEP-2002; 2002US-0411052P.
XX
PR 17-SEP-2002; 2002US-0411055P.
XX
PR 17-SEP-2002; 2002US-0411073P.
XX
PR 17-SEP-2002; 2002US-0411082P.
XX
PR 17-SEP-2002; 2002US-0411101P.
XX
PR 17-SEP-2002; 2002US-0411111P.
XX
PR 18-APR-2003; 2003US-0463700P.
```

```
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471366P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX PA
XX PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX PI Halenbeck RF, Huang MM, Kothakota S, Halahan L, Linnemann T;
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX
XX PS Claim 1; SEQ ID NO 3140; 428bp; English.
XX
XX CC The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPWEB and is not in the specification.
XX
XX SQ Sequence 669 AA;

Query Match          64.3%; Score 18; DB 8; Length 669;
Best Local Similarity 16.7%; Pred. No. 7.3e-05;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY      1 CXXXXXXXXXXC 12
Db      652 CAATTTTAAAC 663

RESULT 14
ADP31215
ID ADP31215 standard; protein; 711 AA.
XX
AC ADP31215;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human secreted protein SEQ ID #1982.
XX
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX cancer; inflammatory; immune; human secreted protein.
```



```
OS Homo sapiens.
XX
XX WO2004035732-A2.
XX
XX 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 29-AUG-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411011P.
XX 17-SEP-2002; 2002US-0411111P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471306P.
XX 19-MAY-2003; 2003US-047136P.
XX 22-MAY-2003; 2003US-0472420P.
XX 22-MAY-2003; 2003US-0472430P.
XX 09-JUN-2003; 2003US-0476609P.
XX 09-JUN-2003; 2003US-0476641P.
XX 08-JUL-2003; 2003US-0485218P.
XX 08-JUL-2003; 2003US-0485233P.
XX 08-JUL-2003; 2003US-0485244P.
XX 08-JUL-2003; 2003US-0485325P.
XX 14-JUL-2003; 2003US-0486446P.
XX 14-JUL-2003; 2003US-0486480P.
XX 15-JUL-2003; 2003US-0486891P.
XX 15-JUL-2003; 2003US-0486960P.
XX 08-AUG-2003; 2003US-0493341P.
PR
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
XX PI Halenbeck RF, Huang MM, Kochakota S, Hsieh L, Linemann T;
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX WPI; 2004-348438/32.
XX
XX PS Claim 1, SEQ ID NO 3213; 428bp; English.
XX
XX CC The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPOMEB and is not in the specification.
XX
XX SQ Sequence 711 AA;
XX
XX Query Match 64.3%; Score 18; DB 8; Length 711;
XX Best Local Similarity 16.7%; Pred. No. 7.4e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
XX QY 1 CXXXXXXXXXXC 12
XX Db 684 CAAAAAATTAC 695
XX
XX RESULT 15
XX ADP31196
XX ID ADP31196 standard, protein; 739 AA.
XX AC ADP31196;
XX XX
XX DT 12-AUG-2004 (first entry)
XX XX
XX DE Human secreted protein SEQ ID #1963.
XX
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX OS Homo sapiens.
XX
XX PN WO2004035732-A2.
XX
XX PD 29-APR-2004.
XX
XX 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
```

PR	17-SEP-2002;	2002US-0410946P.	PR
PR	17-SEP-2002;	2002US-0410948P.	PR
PR	17-SEP-2002;	2002US-0410949P.	PR
PR	17-SEP-2002;	2002US-0410953P.	PR
PR	17-SEP-2002;	2002US-0410957P.	PR
PR	17-SEP-2002;	2002US-0410958P.	PR
PR	17-SEP-2002;	2002US-0410959P.	PR
PR	17-SEP-2002;	2002US-0410960P.	PR
PR	17-SEP-2002;	2002US-0410961P.	PR
PR	17-SEP-2002;	2002US-0410962P.	PR
PR	17-SEP-2002;	2002US-0411019P.	PR
PR	17-SEP-2002;	2002US-0411022P.	PR
PR	17-SEP-2002;	2002US-0411023P.	PR
PR	17-SEP-2002;	2002US-0411024P.	PR
PR	17-SEP-2002;	2002US-0411032P.	PR
PR	17-SEP-2002;	2002US-0411035P.	PR
PR	17-SEP-2002;	2002US-0411037P.	PR
PR	17-SEP-2002;	2002US-0411041P.	PR
PR	17-SEP-2002;	2002US-0411045P.	PR
PR	17-SEP-2002;	2002US-0411046P.	PR
PR	17-SEP-2002;	2002US-0411048P.	PR
PR	17-SEP-2002;	2002US-0411052P.	PR
PR	17-SEP-2002;	2002US-0411055P.	PR
PR	17-SEP-2002;	2002US-0411073P.	PR
PR	17-SEP-2002;	2002US-0411082P.	PR
PR	17-SEP-2002;	2002US-0411111P.	PR
PR	17-SEP-2002;	2002US-0411111P.	PR
PR	18-APR-2003;	2003US-0463700P.	PR
PR	18-APR-2003;	2003US-0463708P.	PR
PR	18-APR-2003;	2003US-0463716P.	PR
PR	18-APR-2003;	2003US-0463732P.	PR
PR	02-MAY-2003;	2003US-0467199P.	PR
PR	02-MAY-2003;	2003US-0467201P.	PR
PR	02-MAY-2003;	2003US-0467203P.	PR
PR	19-MAY-2003;	2003US-0467230P.	PR
PR	19-MAY-2003;	2003US-0471306P.	PR
PR	22-MAY-2003;	2003US-0471336P.	PR
PR	22-MAY-2003;	2003US-0472420P.	PR
PR	09-JUN-2003;	2003US-0476609P.	PR
PR	09-JUN-2003;	2003US-0476641P.	PR
PR	08-JUL-2003;	2003US-0485218P.	PR
PR	08-JUL-2003;	2003US-0485223P.	PR
PR	08-JUL-2003;	2003US-0485224P.	PR
PR	08-JUL-2003;	2003US-0485325P.	PR
PR	14-JUL-2003;	2003US-0486446P.	PR
PR	14-JUL-2003;	2003US-0486480P.	PR
PR	15-JUL-2003;	2003US-0486891P.	PR
PR	15-JUL-2003;	2003US-0486960P.	PR
PR	08-AUG-2003;	2003US-0493341P.	PR
PR	08-AUG-2003;	2003US-0493370P.	PR
PR	08-AUG-2003;	2003US-0493573P.	PR
PR	08-AUG-2003;	2003US-0493577P.	PR
XX			XX
PA	(FIVE-) FIVE PRIME THERAPEUTICS INC.		XX
PI	Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D,		XX
PI	Halenbeck RF, Huang MM, Kothakota S, Haisan L, Linnemann T;		XX
PI	Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;		XX
XX			XX
DR	WPI; 2004-348438/32.		XX
XX			XX
PT	New nucleic acid molecule for diagnosing, preventing or treating diseases		XX
PT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,		XX
PT	genetic, bacterial and viral diseases.		XX
PS	Claim 1; SEQ ID NO 3194; 428bp; English.		XX
CC	The present invention relates to an isolated nucleic acid molecule		CC
CC	encoding a polypeptide which is believed to be cytostatic,		CC
CC	antiinflammatory, immunosuppressive, antibacterial and vitrucial. The		CC
CC	composition and methods are useful for diagnosing, preventing and		CC

CC	treating diseases such as proliferative (e.g. cancer), inflammatory,
CC	immune, metabolic, genetic, bacterial and viral diseases. The present
CC	sequence represents a human secreted protein. The present sequence is
CC	available on WIPOMED and is not in the specification.
XX	
XX	
SQ	Sequence 739 AA;
Query Match	64.3%; Score 18; DB 8; Length 739;
Best Local Similarity	16.7%; Pred. No. 7,4e-05;
Matches	2; Conservative 0; Mismatches 10; Indels 0; Gaps 0
OY	1 CXXXXXXXXXC 12
Db	226 CAAAAAAAAAAC 237
RESULT 16	
ADP31244	
ID	ADP31244 standard; protein; 771 AA.
XX	
AC	ADP31244;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Human secreted protein SEQ ID #2011.
XX	
KW	Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW	cancer; inflammatory; immune; human secreted protein.
XX	
OS	Homo sapiens.
XX	
PN	WO200403732-A2.
XX	
PD	29-APR-2004.
XX	
PE	28-AUG-2003; 2003WO-US026780.
XX	
FR	29-AUG-2002; 2002US-0406576P.
PR	29-AUG-2002; 2002US-0406579P.
PR	29-AUG-2002; 2002US-0406585P.
PR	29-AUG-2002; 2002US-0406588P.
PR	29-AUG-2002; 2002US-0406608P.
PR	29-AUG-2002; 2002US-0406611P.
PR	29-AUG-2002; 2002US-0406612P.
PR	29-AUG-2002; 2002US-0406616P.
PR	29-AUG-2002; 2002US-0406640P.
PR	29-AUG-2002; 2002US-0406642P.
PR	29-AUG-2002; 2002US-0406646P.
PR	29-AUG-2002; 2002US-0406653P.
PR	29-AUG-2002; 2002US-0406655P.
PR	29-AUG-2002; 2002US-0406666P.
PR	17-SEP-2002; 2002US-0410946P.
PR	17-SEP-2002; 2002US-0410947P.
PR	17-SEP-2002; 2002US-0410948P.
PR	17-SEP-2002; 2002US-0410949P.
PR	17-SEP-2002; 2002US-0410953P.
PR	17-SEP-2002; 2002US-0410957P.
PR	17-SEP-2002; 2002US-0410958P.
PR	17-SEP-2002; 2002US-0410959P.
PR	17-SEP-2002; 2002US-0410960P.
PR	17-SEP-2002; 2002US-0410961P.
PR	17-SEP-2002; 2002US-0410962P.
PR	17-SEP-2002; 2002US-0411019P.
PR	17-SEP-2002; 2002US-0411023P.
PR	17-SEP-2002; 2002US-0411023P.
PR	17-SEP-2002; 2002US-0411024P.
PR	17-SEP-2002; 2002US-0411032P.
PR	17-SEP-2002; 2002US-0411035P.
PR	17-SEP-2002; 2002US-0411037P.
PR	17-SEP-2002; 2002US-0411041P.
PR	17-SEP-2002; 2002US-0411045P.
PR	17-SEP-2002; 2002US-0411046P.
PR	17-SEP-2002; 2002US-0411048P.



PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kothakota S, Haislan L, Linemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
DR WPI; 2004-348438/32.  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3218; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 876 AA;  
XX  
Query Match 64.3%; Score 18; DB 8; Length 876;  
Best Local Similarity 16.7%; Pred. No. 7.7e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXC 12  
DB 28 CATATATAAATC 39  
XX  
RESULT 18  
ADP31688  
ID ADP31688 standard; protein; 882 AA.  
XX  
AC ADP31688;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2455.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406665P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463715P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471305P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476619P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
XX PI Halenbeck RF, Huang MM, Kothakota S, Haislan L, Linemann T,  
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.  
PT

XX Claim 1, SEQ ID NO 3686; 428pp; English.  
PS  
XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present  
XX sequence represents a human secreted protein. The present sequence is  
XX available on WIPOWEB and is not in the specification.  
SQ Sequence 882 AA;

PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486466P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halsebeck RF, Huang MM, Kochakota S, Haislan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
PS  
PS Claim 1; SEQ ID NO 3515; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic.  
XX CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The  
XX CC composition and methods are useful for diagnosing, preventing and  
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present  
XX CC sequence represents a human secreted protein. The present sequence is  
XX CC available on WIPOMB and is not in the specification.  
XX  
XX SQ Sequence 1044 AA;  
XX  
XX  
XX Query Match 64.3%; Score 18; DB 8; Length 1044;  
XX Best Local Similarity 16.7%; Pred. No. 8e-05;  
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
OY 1 CXXXXXXXXXXC 12  
Db 325 CATATAATTTC 336  
RESULT 21  
ADP31175  
ID ADP31175 standard; protein; 1086 AA.  
XX  
XX ADP31175;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #1942.  
XX  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
OS Homo sapiens.  
XX  
XX  
XX WO2004035732-A2.  
XX

PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406645P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.  
XX 29-AUG-2002; 2002US-0406666P.  
XX 17-SEP-2002; 2002US-0410946P.  
XX 17-SEP-2002; 2002US-0410947P.  
XX 17-SEP-2002; 2002US-0410948P.  
XX 17-SEP-2002; 2002US-0410949P.  
XX 17-SEP-2002; 2002US-0410953P.  
XX 17-SEP-2002; 2002US-0410957P.  
XX 17-SEP-2002; 2002US-0410958P.  
XX 17-SEP-2002; 2002US-0410959P.  
XX 17-SEP-2002; 2002US-0410960P.  
XX 17-SEP-2002; 2002US-0410961P.  
XX 17-SEP-2002; 2002US-0410962P.  
XX 17-SEP-2002; 2002US-0410963P.  
XX 17-SEP-2002; 2002US-0411019P.  
XX 17-SEP-2002; 2002US-0411022P.  
XX 17-SEP-2002; 2002US-0411023P.  
XX 17-SEP-2002; 2002US-0411024P.  
XX 17-SEP-2002; 2002US-0411032P.  
XX 17-SEP-2002; 2002US-0411035P.  
XX 17-SEP-2002; 2002US-0411037P.  
XX 17-SEP-2002; 2002US-0411041P.  
XX 17-SEP-2002; 2002US-0411045P.  
XX 17-SEP-2002; 2002US-0411046P.  
XX 17-SEP-2002; 2002US-0411048P.  
XX 17-SEP-2002; 2002US-0411052P.  
XX 17-SEP-2002; 2002US-0411055P.  
XX 17-SEP-2002; 2002US-0411073P.  
XX 17-SEP-2002; 2002US-0411082P.  
XX 17-SEP-2002; 2002US-0411101P.  
XX 17-SEP-2002; 2002US-0411111P.  
XX 18-APR-2003; 2003US-0463700P.  
XX 18-APR-2003; 2003US-0463708P.  
XX 18-APR-2003; 2003US-0463716P.  
XX 18-APR-2003; 2003US-0463732P.  
XX 02-MAY-2003; 2003US-0467199P.  
XX 02-MAY-2003; 2003US-0467201P.  
XX 02-MAY-2003; 2003US-0467203P.  
XX 02-MAY-2003; 2003US-0467230P.  
XX 19-MAY-2003; 2003US-0471306P.  
XX 19-MAY-2003; 2003US-0471336P.  
XX 22-MAY-2003; 2003US-0472420P.  
XX 22-MAY-2003; 2003US-0472430P.  
XX 09-JUN-2003; 2003US-0476609P.  
XX 09-JUN-2003; 2003US-0476641P.  
XX 08-JUL-2003; 2003US-0485218P.  
XX 08-JUL-2003; 2003US-0485223P.  
XX 08-JUL-2003; 2003US-0485224P.  
XX 08-JUL-2003; 2003US-0485325P.  
XX 08-JUL-2003; 2003US-0485352P.  
XX 14-JUL-2003; 2003US-0486466P.  
XX 14-JUL-2003; 2003US-0486480P.  
XX 15-JUL-2003; 2003US-0486891P.  
XX 15-JUL-2003; 2003US-0486960P.  
XX 08-AUG-2003; 2003US-0493341P.  
XX 08-AUG-2003; 2003US-0493370P.  
XX 08-AUG-2003; 2003US-0493573P.  
XX 08-AUG-2003; 2003US-0493577P.  
XX



XX Sequence 1113 AA;  
SQ  
Query Match 64.3%; Score 18; DB 8; Length 1113;  
Best Local Similarity 16.7%; Pred. No. 8.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
OY 1 CXXXXXXXXXXC 12  
DB 18 CATAATTTCAC 29  
RESULT 23  
ADA15725  
ID ADA15725 standard; protein; 1128 AA.  
XX  
AC ADA15725;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE C. elegans neuromuscular junction GABA receptor complex subunit #6.  
XX  
KW Nematode;  
XX neuromuscular junction gamma-aminobutyric acid receptor complex;  
XX GABA receptor; parasitic plant pathogen; agricultural industry;  
XX crop protection; soil treatment.  
XX  
OS Caenorhabditis elegans.  
XX  
PN US2003065144-A1.  
XX  
PD 03-APR-2003.  
XX  
PF 24-MAY-2002; 2002US-00156240.  
XX  
PR 09-NOV-1998; 98US-0107127P.  
PR 08-NOV-1999; 99US-00436063.  
XX  
PA (UTAH ) UNITV UTAH RES FOUND.  
XX  
PI Bamber BA, Jorgensen EM;  
XX  
DR WPI; 2003-540802/51.  
DR N-PSDB; ADA15726.  
XX  
PT New nematode neuromuscular junction GABA receptor complex, useful for  
XX crop protection or soil treatment.  
PS Claim 21; Page 54-56; 84pp; English.  
XX  
XX The present invention relates to a nematode neuromuscular junction gamma-  
CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant  
CC pathogens which can cause major damage to crops in the agricultural  
CC industry. The nematode neuromuscular junction GABA receptor complex is  
CC useful for crop protection or soil treatment. The present sequence  
CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor  
CC complex subunit.  
XX  
SQ Sequence 1128 AA;  
Query Match 64.3%; Score 18; DB 6; Length 1128;  
Best Local Similarity 16.7%; Pred. No. 8.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
OY 1 CXXXXXXXXXXC 12  
DB 390 CAAATATTTTC 401  
RESULT 24  
ADP30741  
ID ADP30741 standard; protein; 1134 AA.  
XX

AC ADP30741;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1508.  
XX  
KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406615P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467206P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.



PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485225P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
(FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX PA  
XX PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halembeck RF, Huang MM, Kochakota S, Halsehan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX MPI; 2004-348438/32.  
XX  
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX PT genetic, bacterial and viral diseases.  
XX PS  
XX PS Claim 1; SEQ ID NO 2739; 428bp; English.  
XX  
XX CC The present invention relates to an isolated nucleic acid molecule  
XX CC encoding a polypeptide which is believed to be cytostatic,  
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX CC composition and methods are useful for diagnosing, preventing and  
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present  
XX CC sequence represents a human secreted protein. The present sequence is  
XX CC available on WIPWEB and is not in the specification.  
XX  
XX SQ Sequence 1134 AA;  
XX  
XX Query Match 64.3%; Score 18; DB 8; Length 1134;  
XX Best Local Similarity 16.7%; Pred. No. 8.1e-05;  
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
XX  
XX QY 1 CXXXXXXXXXXC 12  
XX Db 51 CTATATATTATAC 62  
XX  
XX RESULT 25  
XX ADP30924  
XX ID ADP30924 standard; protein; 1134 AA.  
XX AC  
XX ADP30924;  
XX  
XX 12-AUG-2004 (first entry)  
XX  
XX DE Human secreted protein SEQ ID #1691.  
XX  
XX KM Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX KM cancer; inflammatory; immune; human secreted protein.  
XX OS Homo sapiens.  
XX PN WO2004035732-A2.  
XX  
XX PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 29-AUG-2002; 2002US-0406667P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0410963P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467209P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485239P.  
PR 08-JUL-2003; 2003US-0485239P.  
PR 08-JUL-2003; 2003US-0485242P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX PA  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
XX PI Halembeck RF, Huang MM, Kochakota S, Halsehan L, Linnemann T;  
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX MPI; 2004-348438/32.  
XX  
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases

PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 2922; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic.  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 1134 AA;  
  
Query Match 64.3%; Score 18; DB 8; Length 1134;  
Best Local Similarity 16.7%; Pred. No. 8.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXC 12  
|  
DB 322 CTAATAAATAC 333  
  
RESULT 26  
ADP31046  
ID ADP31046 standard; protein; 1168 AA.  
XX  
AC ADP31046;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1813.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX WO2004035732-A2.  
XX  
XX  
XX PD 29-APR-2004.  
XX  
XX PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX PR 29-AUG-2002; 2002US-0406576P.  
XX PR 29-AUG-2002; 2002US-0406579P.  
XX PR 29-AUG-2002; 2002US-0406585P.  
XX PR 29-AUG-2002; 2002US-0406588P.  
XX PR 29-AUG-2002; 2002US-0406608P.  
XX PR 29-AUG-2002; 2002US-0406611P.  
XX PR 29-AUG-2002; 2002US-0406612P.  
XX PR 29-AUG-2002; 2002US-0406616P.  
XX PR 29-AUG-2002; 2002US-0406640P.  
XX PR 29-AUG-2002; 2002US-0406642P.  
XX PR 29-AUG-2002; 2002US-0406646P.  
XX PR 29-AUG-2002; 2002US-0406653P.  
XX PR 29-AUG-2002; 2002US-0406655P.  
XX PR 29-AUG-2002; 2002US-0406666P.  
XX PR 17-SEP-2002; 2002US-0410946P.  
XX PR 17-SEP-2002; 2002US-0410947P.  
XX PR 17-SEP-2002; 2002US-0410948P.  
XX PR 17-SEP-2002; 2002US-0410949P.  
XX PR 17-SEP-2002; 2002US-0410953P.  
XX PR 17-SEP-2002; 2002US-0410957P.  
XX PR 17-SEP-2002; 2002US-0410958P.  
XX PR 17-SEP-2002; 2002US-0410959P.  
XX PR 17-SEP-2002; 2002US-0410960P.  
XX PR 17-SEP-2002; 2002US-0410961P.  
XX PR 17-SEP-2002; 2002US-0410962P.  
XX PR 17-SEP-2002; 2002US-0411019P.  
XX PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-048646P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
XX  
XX PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX PI Williams LT, Chu K, Lee E, Hestir K, Beurang PA, Behrens D,  
PI Halenbeck RF, Huang MW, Kothakota S, Haisan L, Linemann T,  
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX DR WPI; 2004-348438/32.  
XX  
XX PS New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX PT genetic, bacterial and viral diseases.  
XX  
XX PS Claim 1; SEQ ID NO 3044; 428bp; English.  
XX  
XX CC The present invention relates to an isolated nucleic acid molecule  
XX CC encoding a polypeptide which is believed to be cytostatic.  
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX CC composition and methods are useful for diagnosing, preventing and  
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present  
XX CC sequence represents a human secreted protein. The present sequence is  
XX CC available on WIPOMEB and is not in the specification.  
XX  
SQ Sequence 1168 AA;  
  
Query Match 64.3%; Score 18; DB 8; Length 1168;  
Best Local Similarity 16.7%; Pred. No. 8.2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXC 12  
|



PF 28-AUG-2003; 2003WO-US026780.  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-047609P.  
PR 09-JUN-2003; 2003US-047609P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-048525P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493373P.  
PR 08-AUG-2003; 2003US-0493577P.  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX

PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halebek RP, Huang MM, Kotnakota S, Hsiehan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
XX  
DR  
XX  
PT New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
PS Claim 1; SEQ ID NO 3531; 428bp; English.  
XX  
CC The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
SQ Sequence 1260 AA;  
XX  
Query Match 64.3%; Score 18; DB 8; Length 1260;  
Best Local Similarity 16.7%; Pred. No. 8.3e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXXC 12  
Db 757 CAATTAATAATAC 768  
RESULT 29  
ADP30675  
ID ADP30675 standard; protein; 1289 AA.  
XX  
AC ADP30675;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1442.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX  
PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.

PR	17-SEP-2002;	2002US-0410958P.
PR	17-SEP-2002;	2002US-0410959P.
PR	17-SEP-2002;	2002US-0410960P.
PR	17-SEP-2002;	2002US-0410961P.
PR	17-SEP-2002;	2002US-0410962P.
PR	17-SEP-2002;	2002US-0411019P.
PR	17-SEP-2002;	2002US-0411022P.
PR	17-SEP-2002;	2002US-0411023P.
PR	17-SEP-2002;	2002US-0411024P.
PR	17-SEP-2002;	2002US-0411024P.
PR	17-SEP-2002;	2002US-0411046P.
PR	17-SEP-2002;	2002US-0411048P.
PR	17-SEP-2002;	2002US-0411052P.
PR	17-SEP-2002;	2002US-0411037P.
PR	17-SEP-2002;	2002US-0411041P.
PR	17-SEP-2002;	2002US-0411045P.
PR	17-SEP-2002;	2002US-0411046P.
PR	17-SEP-2002;	2002US-0411082P.
PR	17-SEP-2002;	2002US-0411101P.
PR	17-SEP-2002;	2002US-0411111P.
PR	18-APR-2003;	2003US-0463700P.
PR	18-APR-2003;	2003US-0463708P.
PR	18-APR-2003;	2003US-0463716P.
PR	18-APR-2003;	2003US-0463732P.
PR	02-MAY-2003;	2003US-0467199P.
PR	02-MAY-2003;	2003US-0467201P.
PR	02-MAY-2003;	2003US-0467203P.
PR	02-MAY-2003;	2003US-0467230P.
PR	19-MAY-2003;	2003US-0471306P.
PR	19-MAY-2003;	2003US-0471336P.
PR	22-MAY-2003;	2003US-0472420P.
PR	22-MAY-2003;	2003US-0472430P.
PR	09-JUN-2003;	2003US-0476609P.
PR	09-JUN-2003;	2003US-0476641P.
PR	08-JUL-2003;	2003US-0485218P.
PR	08-JUL-2003;	2003US-0485223P.
PR	08-JUL-2003;	2003US-0485224P.
PR	08-JUL-2003;	2003US-0485325P.
PR	14-JUL-2003;	2003US-0486446P.
PR	14-JUL-2003;	2003US-0486480P.
PR	15-JUL-2003;	2003US-0486891P.
PR	15-JUL-2003;	2003US-0486960P.
PR	08-AUG-2003;	2003US-0493341P.
PR	08-AUG-2003;	2003US-0493370P.
PR	08-AUG-2003;	2003US-0493573P.
PR	08-AUG-2003;	2003US-0493577P.
XX	(FIVE-) FIVE PRIME THERAPEUTICS INC.	
XX		
PI	Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;	
PI	Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;	
PI	Plerence K, Wang Y, Wong JGP, Wu G, Zhang H;	
XX		
DR	WPI: 2004-348438/32.	
XX		
FT	New nucleic acid molecule for diagnosing, preventing or treating diseases	
FT	such as proliferative (e.g. cancer), inflammatory, immune, metabolic,	
FT	genetic, bacterial and viral diseases.	
XX		
PS	Claim 1; SEQ ID NO 2673; 428bp; English.	
XX		
CC	The present invention relates to an isolated nucleic acid molecule	
CC	encoding a polypeptide which is believed to be cytostatic,	
CC	antiinflammatory, immunosuppressive, antibacterial and virucidal. The	
CC	composition and methods are useful for diagnosing, preventing and	
CC	treating diseases such as proliferative (e.g. cancer), inflammatory,	
CC	immune, metabolic, genetic, bacterial and viral diseases. The present	
CC	sequence represents a human secreted protein. The present sequence is	
CC	available on WIPower and is not in the specification.	
XX		
XX	Sequence 1289 AA:	

Query Match	64.3%	Score 18;	DB 8;	Length 1289;
Best Local Similarity	16.7%	Pred. No. 8.3e-05;		
Matches 2; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0
QY	1	XXXXXXXXXXC	12	
DB	1076	CATATATAATTC	1087	
RESULT 30				
ID	ADP31357	standard; protein; 1437 AA.		
XX	ADP31357;			
XX				
DT	12-AUG-2004	(first entry)		
XX				
DE	Human secreted protein SEQ ID #2124.			
XX				
KW	Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide; cancer; Inflammatory; Immune; human secreted protein.			
XX				
OS	Homo sapiens.			
XX				
PN	MO2004035732-A2.			
XX				
PD	29-APR-2004.			
XX				
PF	28-AUG-2003; 2003WO-US026780.			
XX				
PR	29-AUG-2002; 2002US-0406576P.			
PR	29-AUG-2002; 2002US-0406579P.			
PR	29-AUG-2002; 2002US-0406585P.			
PR	29-AUG-2002; 2002US-0406588P.			
PR	29-AUG-2002; 2002US-0406608P.			
PR	29-AUG-2002; 2002US-0406611P.			
PR	29-AUG-2002; 2002US-0406612P.			
PR	29-AUG-2002; 2002US-0406616P.			
PR	29-AUG-2002; 2002US-0406640P.			
PR	29-AUG-2002; 2002US-0406642P.			
PR	29-AUG-2002; 2002US-0406646P.			
PR	29-AUG-2002; 2002US-0406653P.			
PR	29-AUG-2002; 2002US-0406655P.			
PR	29-AUG-2002; 2002US-0406666P.			
PR	17-SEP-2002; 2002US-0410946P.			
PR	17-SEP-2002; 2002US-0410947P.			
PR	17-SEP-2002; 2002US-0410948P.			
PR	17-SEP-2002; 2002US-0410949P.			
PR	17-SEP-2002; 2002US-0410953P.			
PR	17-SEP-2002; 2002US-0410957P.			
PR	17-SEP-2002; 2002US-0410958P.			
PR	17-SEP-2002; 2002US-0410959P.			
PR	17-SEP-2002; 2002US-0410960P.			
PR	17-SEP-2002; 2002US-0410961P.			
PR	17-SEP-2002; 2002US-0410962P.			
PR	17-SEP-2002; 2002US-0411019P.			
PR	17-SEP-2002; 2002US-0411022P.			
PR	17-SEP-2002; 2002US-0411023P.			
PR	17-SEP-2002; 2002US-0411024P.			
PR	17-SEP-2002; 2002US-0411032P.			
PR	17-SEP-2002; 2002US-0411033P.			
PR	17-SEP-2002; 2002US-0411037P.			
PR	17-SEP-2002; 2002US-0411041P.			
PR	17-SEP-2002; 2002US-0411045P.			
PR	17-SEP-2002; 2002US-0411046P.			
PR	17-SEP-2002; 2002US-0411048P.			
PR	17-SEP-2002; 2002US-0411052P.			
PR	17-SEP-2002; 2002US-0411055P.			
PR	17-SEP-2002; 2002US-0411073P.			
PR	17-SEP-2002; 2002US-0411082P.			
PR	17-SEP-2002; 2002US-0411101P.			
PR	17-SEP-2002; 2002US-0411111P.			



PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX MPI, 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 3175; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.  
XX  
XX Sequence 1454 AA;  
SQ  
Query Match 64.3%; Score 18; DB 8; Length 1454;  
Best Local Similarity 16.7%; Pred. No. 8.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXXC 12  
DB 1094 CATTATATATAC 1105  
RESULT 32  
ADP30557  
ID ADP30557 standard; protein; 1480 AA.  
XX  
AC ADP30557;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
XX Human secreted protein SEQ ID #1324.  
DE  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX MO2004035732-A2.  
XX  
XX PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
XX 29-AUG-2002; 2002US-0406579P.  
XX 29-AUG-2002; 2002US-0406585P.  
XX 29-AUG-2002; 2002US-0406588P.  
XX 29-AUG-2002; 2002US-0406608P.  
XX 29-AUG-2002; 2002US-0406611P.  
XX 29-AUG-2002; 2002US-0406612P.  
XX 29-AUG-2002; 2002US-0406616P.  
XX 29-AUG-2002; 2002US-0406640P.  
XX 29-AUG-2002; 2002US-0406642P.  
XX 29-AUG-2002; 2002US-0406646P.  
XX 29-AUG-2002; 2002US-0406653P.  
XX 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485252P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486896P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX MPI, 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2555; 428bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The

CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.

CC Sequence 1480 AA;

Query Match 64.3%; Score 18; DB 8; Length 1480;  
Best Local Similarity 16.7%; Pred. No. 8.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 12

Db 1286 CTPAAATATTAAC 1297

#### RESULT 33

ABB09437 standard; protein; 1588 AA.

AC ABB09437;

DT 01-JUL-2002 (first entry)

DE H. influenzae DXR related polypeptide sequence.

KW DXR; reductoisomerase; enzyme; non-mevalonate isoprenoid; menaquinone;  
KW ubiquinone; vitamin; ear infection; conjunctivitis; meningitis;  
KW pneumonia; conjunctivitis; bacteraemia; sinusitis; pleural empyema;  
KW endocarditis; epiglottitis.

OS Haemophilus influenzae.

Key Location/Qualifiers  
Region 241..1431

/notes="region that appears to be accidentally inserted  
into the sequence, consisting the DXR encoding DNA  
sequence represented as an amino acid sequence in three  
letter code"

WO200211673-A2.

14-FEB-2002.

09-AUG-2001; 2001WO-US024950.

09-AUG-2000; 2000US-0223809P.

(SMK ) SMITHKLINE BEECHAM CORP.  
(SMK ) SMITHKLINE BEECHAM PLC.

Jaworski DD, Payne DJ, Slater-Radoski CE, Yan K;

WPI; 2002-241698/29.

Modulating Haemophilus influenzae DXR reductoisomerase enzyme activity,  
useful for treating mammals or tissues infected with H. influenzae (e.g.  
ear infections or pneumonia) by contacting the enzyme with a modulator of  
its activity.

Disclosure; Page 40-44; 44pp; English.

The invention relates to modulating an activity of a DXR reductoisomerase  
enzyme of Haemophilus influenzae, comprising contacting the enzyme with a  
compound that modulates non-mevalonate isoprenoid biosynthesis -  
synthesis of menaquinone or ubiquinone. Compounds of the invention act as  
virocidases. The method is useful for treating a mammal or mammalian tissue  
infected with H. influenzae having DXR reductoisomerase enzyme, e.g. a  
human or a domestic animal. In particular, the method is useful for  
treating ear infections, conjunctivitis, meningitis, pneumonia,  
conjunctivitis, bacteraemia, sinusitis, pleural empyema, endocarditis and  
epiglottitis. The current sequence represents a H. influenzae DXR

CC reductoisomerase enzyme related polypeptide sequence. Note: The current  
CC sequence contains within it the amino acid sequence given in record  
CC ABB09436 (DXR enzyme), but this is broken up by a large insertion that  
CC appears to be accidentally inserted into the sequence, consisting the DXR  
CC encoding DNA sequence represented as an amino acid sequence in three  
CC letter code

CC Sequence 1588 AA;

Query Match 64.3%; Score 18; DB 5; Length 1588;  
Best Local Similarity 16.7%; Pred. No. 8.7e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXC 12

Db 431 CTPAAATTTTAC 442

#### RESULT 34

ADP30660 standard; protein; 1617 AA.

AC ADP30660;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1427.

KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

WO2004035732-A2.

29-APR-2004.

28-AUG-2003; 2003WO-US026780.

29-AUG-2002; 2002US-0406576P.

29-AUG-2002; 2002US-0406579P.

29-AUG-2002; 2002US-0406585P.

29-AUG-2002; 2002US-0406588P.

29-AUG-2002; 2002US-0406608P.

29-AUG-2002; 2002US-0406611P.

29-AUG-2002; 2002US-0406612P.

29-AUG-2002; 2002US-0406616P.

29-AUG-2002; 2002US-0406640P.

29-AUG-2002; 2002US-0406642P.

29-AUG-2002; 2002US-0406646P.

29-AUG-2002; 2002US-0406651P.

29-AUG-2002; 2002US-0406655P.

29-AUG-2002; 2002US-0406666P.

17-SEP-2002; 2002US-0410946P.

17-SEP-2002; 2002US-0410947P.

17-SEP-2002; 2002US-0410948P.

17-SEP-2002; 2002US-0410949P.

17-SEP-2002; 2002US-0410953P.

17-SEP-2002; 2002US-0410957P.

17-SEP-2002; 2002US-0410958P.

17-SEP-2002; 2002US-0410959P.

17-SEP-2002; 2002US-0410960P.

17-SEP-2002; 2002US-0410961P.

17-SEP-2002; 2002US-0410962P.

17-SEP-2002; 2002US-0411019P.

17-SEP-2002; 2002US-0411022P.

17-SEP-2002; 2002US-0411023P.

17-SEP-2002; 2002US-0411024P.

17-SEP-2002; 2002US-0411032P.

17-SEP-2002; 2002US-0411035P.

17-SEP-2002; 2002US-0411037P.

17-SEP-2002; 2002US-0411041P.

17-SEP-2002; 2002US-0411045P.



PR 17-SEP-2002; 2002US-0411046P.  
 PR 17-SEP-2002; 2002US-0411048P.  
 PR 17-SEP-2002; 2002US-0411052P.  
 PR 17-SEP-2002; 2002US-0411055P.  
 PR 17-SEP-2002; 2002US-0411073P.  
 PR 17-SEP-2002; 2002US-0411082P.  
 PR 17-SEP-2002; 2002US-0411011P.  
 PR 17-SEP-2002; 2002US-0411111P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463708P.  
 PR 18-APR-2003; 2003US-0463716P.  
 PR 18-APR-2003; 2003US-0463732P.  
 PR 02-MAY-2003; 2003US-0467199P.  
 PR 02-MAY-2003; 2003US-0467201P.  
 PR 02-MAY-2003; 2003US-0467203P.  
 PR 02-MAY-2003; 2003US-0467230P.  
 PR 19-MAY-2003; 2003US-0471306P.  
 PR 19-MAY-2003; 2003US-0471336P.  
 PR 22-MAY-2003; 2003US-0472420P.  
 PR 22-MAY-2003; 2003US-0472430P.  
 PR 09-JUN-2003; 2003US-0476609P.  
 PR 09-JUN-2003; 2003US-0476641P.  
 PR 08-JUL-2003; 2003US-0485218P.  
 PR 08-JUL-2003; 2003US-0485223P.  
 PR 08-JUL-2003; 2003US-0485224P.  
 PR 08-JUL-2003; 2003US-0485325P.  
 PR 14-JUL-2003; 2003US-0486446P.  
 PR 14-JUL-2003; 2003US-0486480P.  
 PR 15-JUL-2003; 2003US-0486891P.  
 PR 15-JUL-2003; 2003US-0486960P.  
 PR 08-AUG-2003; 2003US-0493341P.  
 PR 08-AUG-2003; 2003US-0493370P.  
 PR 08-AUG-2003; 2003US-0493573P.  
 PR 08-AUG-2003; 2003US-0493577P.

# (FIVE-) FIVE PRIME THERAPEUTICS INC.

XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
 PI Halenbeck RF, Huang MM, Kochakota S, Haishan L, Linnemann T;  
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
 XX MPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
 PT genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 2658; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
 CC encoding a polypeptide which is believed to be cytostatic,  
 CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
 CC composition and methods are useful for diagnosing, preventing and  
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
 CC immune, metabolic, genetic, bacterial and viral diseases. The present  
 CC sequence represents a human secreted protein. The present sequence is  
 CC available on WIPOMB and is not in the specification.

XX Sequence 1617 AA;

QY Query Match 64.3%; Score 18; DB 8; Length 1617;  
 Best Local Similarity 16.7%; Pred. No. 8.7e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
 Db 71 CTATATTAATTC 82

RESULT 35  
 ADA15715  
 ID ADA15715 standard; protein; 1652 AA.  
 XX

AC ADA15715;

XX 06-NOV-2003 (first entry)

XX C. elegans neuromuscular junction GABA receptor complex subunit #1.

XX Nematode;

XX neuromuscular junction gamma-aminobutyric acid receptor complex;

XX GABA receptor; parasitic plant pathogen; agricultural industry;

XX crop protection; soil treatment.

XX Caenorhabditis elegans.

XX US2003065144-A1.

XX 03-APR-2003.

XX 24-MAY-2002; 2002US-00156240.

XX 09-NOV-1998; 98US-0107727P.

XX 08-NOV-1999; 99US-00436063.

XX (UTAH) UNIV UTAH RES FOUND.

XX Bamber BA, Jorgensen EM;

XX MPI; 2003-540802/51.

XX New nematode neuromuscular junction GABA receptor complex, useful for

XX crop protection or soil treatment.

XX Claim 21; Page 20-24; 84p; English.

XX The present invention relates to a nematode neuromuscular junction gamma-

XX aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant

XX pathogens which can cause major damage to crops in the agricultural

XX industry. The nematode neuromuscular junction GABA receptor complex is

XX useful for crop protection or soil treatment. The present sequence

XX represents a Caenorhabditis elegans neuromuscular junction GABA receptor

XX complex subunit.

XX Sequence 1652 AA;

QY Query Match 64.3%; Score 18; DB 6; Length 1652;  
 Best Local Similarity 16.7%; Pred. No. 8.8e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
 Db 891 CAATATTTTTC 902

ADP30654

ADP30654 standard; protein; 1725 AA.

ADP30654;

12-AUG-2004 (first entry)

Human secreted protein SEQ ID #1421.

Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

cancer; inflammatory; immune; human secreted protein.

Homo sapiens.

WO2004035732-A2.

29-APR-2004.

28-AUG-2003; 2003WO-US026780.

PR 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 19-MAY-2003; 2003US-047136P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Halshen L, Linnemann T;

PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.  
DR  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2652; 428pp; English.  
PS  
XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
XX available on WIPWEB and is not in the specification.  
SQ Sequence 1725 AA;  
XX  
SQ  
Query Match 64.3%; Score 18; DB 8; Length 1725;  
Best Local Similarity 16.7%; Pred. No. 8.9e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXXC 12  
DB 828 CTATTTTAAAC 839  
RESULT 37  
ADP30642  
ID ADP30642 standard; protein; 1833 AA.  
XX  
XX ADP30642;  
AC  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #1409.  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.  
XX  
XX  
XX WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 26-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 29-AUG-2002; 2002US-0410946P.  
PR 29-AUG-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.

Best Local Similarity 16.7%; Pred. No. 9e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXX 12  
Db 462 CTATTTTAAAC 473

## RESULT 38

ADP31178

ID ADP31178 standard; protein; 2088 AA.

XX ADP31178;

DT 12-AUG-2004 (first entry)

DE Human secreted protein SEQ ID #1945.

XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;

KW cancer; inflammatory; immune; human secreted protein.

OS Homo sapiens.

PN WO2004035732-A2.

PD 29-APR-2004.

PF 28-AUG-2003; 2003WO-US026780.

XX 29-AUG-2002; 2002US-0406576P.

PR 29-AUG-2002; 2002US-0406579P.

PR 29-AUG-2002; 2002US-0406585P.

PR 29-AUG-2002; 2002US-0406588P.

PR 29-AUG-2002; 2002US-0406608P.

PR 29-AUG-2002; 2002US-0406611P.

PR 29-AUG-2002; 2002US-0406612P.

PR 29-AUG-2002; 2002US-0406616P.

PR 29-AUG-2002; 2002US-0406642P.

PR 29-AUG-2002; 2002US-0406646P.

PR 29-AUG-2002; 2002US-0406653P.

PR 29-AUG-2002; 2002US-0406655P.

PR 29-AUG-2002; 2002US-0406666P.

PR 17-SEP-2002; 2002US-0410946P.

PR 17-SEP-2002; 2002US-0410947P.

PR 17-SEP-2002; 2002US-0410948P.

PR 17-SEP-2002; 2002US-0410949P.

PR 17-SEP-2002; 2002US-0410953P.

PR 17-SEP-2002; 2002US-0410958P.

PR 17-SEP-2002; 2002US-0410959P.

PR 17-SEP-2002; 2002US-0410960P.

PR 17-SEP-2002; 2002US-0410961P.

PR 17-SEP-2002; 2002US-0410962P.

PR 17-SEP-2002; 2002US-0411019P.

PR 17-SEP-2002; 2002US-0411022P.

PR 17-SEP-2002; 2002US-0411023P.

PR 17-SEP-2002; 2002US-0411024P.

PR 17-SEP-2002; 2002US-0411032P.

PR 17-SEP-2002; 2002US-0411033P.

PR 17-SEP-2002; 2002US-0411035P.

PR 17-SEP-2002; 2002US-0411037P.

PR 17-SEP-2002; 2002US-0411041P.

PR 17-SEP-2002; 2002US-0411045P.

PR 17-SEP-2002; 2002US-0411048P.

PR 17-SEP-2002; 2002US-0411052P.

PR 17-SEP-2002; 2002US-0411055P.

PR 17-SEP-2002; 2002US-0411073P.

PR 17-SEP-2002; 2002US-0411083P.

PR 17-SEP-2002; 2002US-0411101P.

PR 17-SEP-2002; 2002US-0411111P.

PR 18-APR-2003; 2003US-0463700P.

PR 18-APR-2003; 2003US-0463708P.

PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485233P.  
PR 08-JUL-2003; 2003US-0485244P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Hestir K, Beaurang PY, Behrens D;  
PI Halenbeck RF, Huang MM, Kothakota S, Halaban L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX Claim 1; SEQ ID NO 2640; 428pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
XX Sequence 1833 AA;  
SQ

Query Match 64.3%; Score 18; DB 8; Length 1833;

PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485252P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halenbeck RF, Huang MM, Kochakota S, Haisan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.  
XX  
PR New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1; SEQ ID NO 3176; 428pp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOWEB and is not in the specification.  
XX  
SQ Sequence 2088 AA;  
Query Match 64.3%; Score 18; DB 8; Length 2088;  
Best Local Similarity 16.7%; Pred. No. 9.2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXXC 12  
Db 1094 CATTATATATAC 1105  
  
RESULT 39  
ADP31327  
ID ADP31327 standard; protein; 2127 AA.  
XX  
XX ADP31327;  
AC  
XX  
XX 12-AUG-2004 (first entry)  
DT  
XX  
XX Human secreted protein SEQ ID #2094.  
DE  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
OS

XX  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
XX 28-AUG-2003; 2003WO-US026780.  
PF  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411015P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 02-MAY-2003; 2003US-0467236P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.

```
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
XX Williams LT, Chu K, Lee E, Heerir K, Beaurang PA, Behrens D;
PI Halanbeck RF, Huang MM, Kochakota S, Haislan L, Linemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PR genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 3325; 428bp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytostatic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPWEB and is not in the specification.
XX
XX Sequence 2127 AA;
SQ
XX
XX Query Match 64.3%; Score 18; DB 8; Length 2127;
XX Best Local Similarity 16.7%; Pred. No. 9.2e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXXC 12
DB 1464 CTTATTAAAAAC 1475
XX
XX RESULT 40
XX ADP66690
XX ID ADP66690 standard; protein; 2484 AA.
XX
XX ADP66690;
XX
XX 26-AUG-2004 (first entry)
XX
XX Human mismatch repair protein MLH1.
XX
XX PMS1; PMS2; PMSR3; PMSR2; PMSR6; MLH1; GTBP; MSH3; MSH1; PMSR;
XX immunoglobulin; mismatch repair protein; human.
XX
XX Homo sapiens.
XX
XX WO2004046330-A2.
XX
XX 03-JUN-2004.
XX
XX 14-NOV-2003; 2003WO-US036702.
XX
XX 15-NOV-2002; 2002US-0427165P.
XX
XX 10-SEP-2003; 2003US-0501650P.
XX
XX (MORP-) MORPHOTEK INC.
XX
XX Graessl L, Liang S, Nicolaides NE, Saas PM;
XX WPI; 2004-440979/41.
XX
XX Producing mammalian expression cells producing high-affinity antibodies
XX from immunized immunoglobulin-producing cells, by combining cells with
XX antigen, forming parental hybridoma cells and hypermutated hybridoma
XX cells, using myeloma cells.
XX
XX Claim 135; SEQ ID NO 12; 213bp; English.
```

```
XX
XX The invention relates to producing mammalian expression cells e.g.
XX hybridoma cells producing high-affinity and high titer antibodies from in
XX vitro immunized immunoglobulin-producing cells. The method involves:
CC combining donor cells comprising immunoglobulin (Ig)-producing cells with
CC an immunogenic antigen in vitro, fusing the Ig-producing cells with
CC myeloma cells to form parental hybridoma cells, where the hybridoma cells
CC express a dominant negative allele of a mismatch repair gene, incubating
CC the parental hybridoma cells to allow for mutagenesis, thus forming
CC hypermutated hybridoma cells, performing a screen for binding of
CC antibodies to antigen for antibodies produced from the hypermutated
CC hybridoma cells, and selecting hypermutated hybridoma cells that produce
CC antibodies with greater affinity for the antigen than antibodies produced
CC by the parental hybridoma cells, thus producing hybridoma cells producing
CC high-affinity antibodies. In the method, the dominant negative allele of
CC a mismatch repair gene comprises a dominant negative allele of a gene
CC chosen from PMS2, PMS1, PMSR3, PMSR2, PMSR6, MLH1, GTBP, MSH3, MSH1, or
CC MSH1, and homologues of PMSR genes. The method is useful for producing
CC mammalian expression cells e.g., hybridoma cells producing high-affinity
CC and high titer antibodies from in vitro immunized immunoglobulin-
CC producing cells. The present sequence represents a human mismatch repair
CC protein MLH1.
XX
XX Sequence 2484 AA;
SQ
XX
XX Query Match 64.3%; Score 18; DB 8; Length 2484;
XX Best Local Similarity 16.7%; Pred. No. 9.6e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 CXXXXXXXXXXC 12
DB 515 CTTTAAAAAATC 526
XX
XX RESULT 41
XX ADA15721
XX ID ADA15721 standard; protein; 2508 AA.
XX
XX ADA15721;
XX
XX 06-NOV-2003 (first entry)
XX
XX C. elegans neuromuscular junction GABA receptor complex subunit #4.
XX
XX Nematode;
XX neuromuscular junction gamma-aminobutyric acid receptor complex;
XX GABA receptor; parasitic plant pathogen; agricultural industry;
XX crop protection; soil treatment.
XX
XX Caenorhabditis elegans.
XX
XX US2003065144-A1.
XX
XX 03-APR-2003.
XX
XX 24-MAY-2002; 2002US-00156240.
XX
XX 09-NOV-1998; 98US-0107727P.
XX
XX 08-NOV-1999; 99US-00436063.
XX
XX (UTAH ) UNIV UTAH RES FOUND.
XX
XX Bamber BA, Jorgensen EM;
XX WPI; 2003-540802/51.
XX
XX N-PSDB; ADA15722.
XX
XX New nematode neuromuscular junction GABA receptor complex, useful for
XX crop protection or soil treatment.
XX
XX Claim 21; Page 38-44; 84bp; English.
XX
XX The present invention relates to a nematode neuromuscular junction gamma-
```

CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant  
CC pathogens which can cause major damage to crops in the agricultural  
CC industry. The nematode neuromuscular junction GABA receptor complex is  
CC useful for crop protection or soil treatment. The present sequence  
CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor  
CC complex subunit.  
XX  
SQ Sequence 2508 AA;  
  
Query Match 64.3%; Score 18; DB 6; Length 2508;  
Best Local Similarity 16.7%; Pred. No. 9.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 1775 CAAATATTTTTC 1786  
  
RESULT 42  
ADA15717  
ID ADA15717 standard; protein; 2544 AA.  
XX  
AC ADA15717;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE C. elegans neuromuscular junction GABA receptor complex subunit #2.  
XX  
KW Nematode;  
KW neuromuscular junction gamma-aminobutyric acid receptor complex;  
KW GABA receptor; parasitic plant pathogen; agricultural industry;  
KW crop protection; soil treatment.  
XX  
OS Caenorhabditis elegans.  
XX  
PN US2003065144-A1.  
XX  
PD 03-APR-2003.  
XX  
PF 24-MAY-2002; 2002US-00156240.  
XX  
PR 09-NOV-1998; 98US-0107727P.  
PR 08-NOV-1999; 99US-00436063.  
XX  
PA (UTAH ) UNIV UTAH RES FOUND.  
XX  
PI Bamber BA, Jorgensen EM;  
XX  
DR WPI, 2003-540802/51.  
DR N-PSDB; ADA15718.  
XX  
PT New nematode neuromuscular junction GABA receptor complex, useful for  
PT crop protection or soil treatment.  
XX  
PS Claim 21; Page 25-31; 84pp; English.  
XX  
CC The present invention relates to a nematode neuromuscular junction gamma-  
CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant  
CC pathogens which can cause major damage to crops in the agricultural  
CC industry. The nematode neuromuscular junction GABA receptor complex is  
CC useful for crop protection or soil treatment. The present sequence  
CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor  
CC complex subunit.  
XX  
SQ Sequence 2544 AA;  
  
Query Match 64.3%; Score 18; DB 6; Length 2544;  
Best Local Similarity 16.7%; Pred. No. 9.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 1811 CAAATATTTTTC 1822

RESULT 43  
ADA15723  
ID ADA15723 standard; protein; 2601 AA.  
XX  
AC ADA15723;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE C. elegans neuromuscular junction GABA receptor complex subunit #5.  
XX  
KW Nematode;  
KW neuromuscular junction gamma-aminobutyric acid receptor complex;  
KW GABA receptor; parasitic plant pathogen; agricultural industry;  
KW crop protection; soil treatment.  
XX  
OS Caenorhabditis elegans.  
XX  
PN US2003065144-A1.  
XX  
PD 03-APR-2003.  
XX  
PF 24-MAY-2002; 2002US-00156240.  
XX  
PR 09-NOV-1998; 98US-0107727P.  
PR 08-NOV-1999; 99US-00436063.  
XX  
PA (UTAH ) UNIV UTAH RES FOUND.  
XX  
PI Bamber BA, Jorgensen EM;  
XX  
DR WPI, 2003-540802/51.  
DR N-PSDB; ADA15724.  
XX  
PT New nematode neuromuscular junction GABA receptor complex, useful for  
PT crop protection or soil treatment.  
XX  
PS Claim 21; Page 46-52; 84pp; English.  
XX  
CC The present invention relates to a nematode neuromuscular junction gamma-  
CC aminobutyric acid (GABA) receptor complex. Nematodes are parasitic plant  
CC pathogens which can cause major damage to crops in the agricultural  
CC industry. The nematode neuromuscular junction GABA receptor complex is  
CC useful for crop protection or soil treatment. The present sequence  
CC represents a Caenorhabditis elegans neuromuscular junction GABA receptor  
CC complex subunit.  
XX  
SQ Sequence 2601 AA;  
  
Query Match 64.3%; Score 18; DB 6; Length 2601;  
Best Local Similarity 16.7%; Pred. No. 9.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 1251 CAATAATTAATTC 1262  
  
RESULT 44  
ADP31299  
ID ADP31299 standard; protein; 2833 AA.  
XX  
AC ADP31299;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human secreted protein SEQ ID #2066.  
XX  
KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
OS Homo sapiens.

```
XX WO2004035732-A2.
PN
XX
XX
PD 29-APR-2004.
PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
XX 17-SEP-2002; 2002US-0410947P.
XX 17-SEP-2002; 2002US-0410948P.
XX 17-SEP-2002; 2002US-0410949P.
XX 17-SEP-2002; 2002US-0410953P.
XX 17-SEP-2002; 2002US-0410957P.
XX 17-SEP-2002; 2002US-0410958P.
XX 17-SEP-2002; 2002US-0410959P.
XX 17-SEP-2002; 2002US-0410960P.
XX 17-SEP-2002; 2002US-0410961P.
XX 17-SEP-2002; 2002US-0410962P.
XX 17-SEP-2002; 2002US-0411019P.
XX 17-SEP-2002; 2002US-0411022P.
XX 17-SEP-2002; 2002US-0411023P.
XX 17-SEP-2002; 2002US-0411024P.
XX 17-SEP-2002; 2002US-0411032P.
XX 17-SEP-2002; 2002US-0411035P.
XX 17-SEP-2002; 2002US-0411037P.
XX 17-SEP-2002; 2002US-0411041P.
XX 17-SEP-2002; 2002US-0411045P.
XX 17-SEP-2002; 2002US-0411046P.
XX 17-SEP-2002; 2002US-0411048P.
XX 17-SEP-2002; 2002US-0411052P.
XX 17-SEP-2002; 2002US-0411055P.
XX 17-SEP-2002; 2002US-0411073P.
XX 17-SEP-2002; 2002US-0411082P.
XX 17-SEP-2002; 2002US-0411101P.
XX 18-APR-2003; 2003US-0463700P.
XX 18-APR-2003; 2003US-0463708P.
XX 18-APR-2003; 2003US-0463716P.
XX 18-APR-2003; 2003US-0463732P.
XX 02-MAY-2003; 2003US-0467199P.
XX 02-MAY-2003; 2003US-0467201P.
XX 02-MAY-2003; 2003US-0467203P.
XX 02-MAY-2003; 2003US-0467230P.
XX 19-MAY-2003; 2003US-0471336P.
XX 22-MAY-2003; 2003US-0472420P.
XX 22-MAY-2003; 2003US-0472430P.
XX 09-JUN-2003; 2003US-0476609P.
XX 09-JUN-2003; 2003US-0476641P.
XX 08-JUL-2003; 2003US-048518P.
XX 08-JUL-2003; 2003US-0485223P.
XX 08-JUL-2003; 2003US-0485224P.
XX 08-JUL-2003; 2003US-0485325P.
XX 14-JUL-2003; 2003US-0486446P.
XX 14-JUL-2003; 2003US-0486480P.
XX 15-JUL-2003; 2003US-0486891P.
XX 08-AUG-2003; 2003US-0493341P.
XX 08-AUG-2003; 2003US-0493370P.

PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX PA
XX PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,
XX PI Halenbeck RF, Huang PM, Kothakota S, Halsehan L, Linnemann T;
XX PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX DR WPI; 2004-348438/32.
XX
XX PT New nucleic acid molecule for diagnosing, preventing or treating diseases
XX PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
XX PT genetic, bacterial and viral diseases.
XX
XX PS Claim 1, SEQ ID NO 3297; 428bp; English.
XX
XX CC The present invention relates to an isolated nucleic acid molecule
XX CC encoding a polypeptide which is believed to be cytostatic,
XX CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
XX CC composition and methods are useful for diagnosing, preventing and
XX CC treating diseases such as proliferative (e.g. cancer), inflammatory,
XX CC immune, metabolic, genetic, bacterial and viral diseases. The present
XX CC sequence represents a human secreted protein. The present sequence is
XX CC available on WIPOMERB and is not in the specification.
XX
XX SQ Sequence 2833 AA;
XX
XX Query Match 64.3%; Score 18; DB 8; Length 2833;
XX Best Local Similarity 16.7%; Pred. No. 9.8e-05;
XX Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
XX QY 1 CXXXXXXXXXXC 12
XX
XX DB 1876 CATTAATTATATC 1887
XX
XX RESULT 45
XX ADP30572
XX ID ADP30572 standard; protein; 2835 AA.
XX
XX AC ADP30572;
XX
XX XX 12-AUG-2004 (first entry)
XX
XX DE Human secreted protein SEQ ID #1339.
XX
XX KW Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
XX KW cancer; inflammatory; immune; human secreted protein.
XX
XX OS Homo sapiens.
XX
XX PN WO2004035732-A2.
XX
XX PD 29-APR-2004.
XX
XX PF 28-AUG-2003; 2003WO-US026780.
XX
XX 29-AUG-2002; 2002US-0406576P.
XX 29-AUG-2002; 2002US-0406579P.
XX 29-AUG-2002; 2002US-0406585P.
XX 29-AUG-2002; 2002US-0406588P.
XX 29-AUG-2002; 2002US-0406608P.
XX 29-AUG-2002; 2002US-0406611P.
XX 29-AUG-2002; 2002US-0406612P.
XX 29-AUG-2002; 2002US-0406616P.
XX 29-AUG-2002; 2002US-0406640P.
XX 29-AUG-2002; 2002US-0406642P.
XX 29-AUG-2002; 2002US-0406646P.
XX 29-AUG-2002; 2002US-0406653P.
XX 29-AUG-2002; 2002US-0406655P.
XX 29-AUG-2002; 2002US-0406666P.
XX 17-SEP-2002; 2002US-0410946P.
```

PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411019P.  
PR 17-SEP-2002; 2002US-0411022P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.  
PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485252P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486896P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
PR XX  
PR XX  
PA (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PI Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D,  
PI Halembeck RF, Huang MM, Kotnakota S, Haishan L, Linnemann T,  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX MPI; 2004-348438/32.  
XX  
XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.  
XX  
XX  
PS Claim 1: SEQ ID NO 2570; 426bp; English.  
XX  
XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,

CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPWEB and is not in the specification.  
XX  
SQ Sequence 2835 AA;  
Query Match 64.3%; Score 18; DB 8; Length 2835;  
Best Local Similarity 16.7%; Pred. No. 9.8e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXXC 12  
Db 884 CAAAAATTATTC 895  
RESULT 46  
ADP30667  
ID ADP30667 standard; protein; 3411 AA.  
XX  
AC ADP30667;  
XX  
XX 12-AUG-2004 (first entry)  
DT XX  
XX  
DE Human secreted protein SEQ ID #1434.  
XX  
XX  
XX Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;  
KW cancer; inflammatory; immune; human secreted protein.  
XX  
XX Homo sapiens.  
PN WO2004035732-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 28-AUG-2003; 2003WO-US026780.  
XX  
XX 29-AUG-2002; 2002US-0406576P.  
PR 29-AUG-2002; 2002US-0406579P.  
PR 29-AUG-2002; 2002US-0406585P.  
PR 29-AUG-2002; 2002US-0406588P.  
PR 29-AUG-2002; 2002US-0406608P.  
PR 29-AUG-2002; 2002US-0406611P.  
PR 29-AUG-2002; 2002US-0406612P.  
PR 29-AUG-2002; 2002US-0406616P.  
PR 29-AUG-2002; 2002US-0406640P.  
PR 29-AUG-2002; 2002US-0406642P.  
PR 29-AUG-2002; 2002US-0406646P.  
PR 29-AUG-2002; 2002US-0406653P.  
PR 29-AUG-2002; 2002US-0406655P.  
PR 29-AUG-2002; 2002US-0406666P.  
PR 17-SEP-2002; 2002US-0410946P.  
PR 17-SEP-2002; 2002US-0410947P.  
PR 17-SEP-2002; 2002US-0410948P.  
PR 17-SEP-2002; 2002US-0410949P.  
PR 17-SEP-2002; 2002US-0410953P.  
PR 17-SEP-2002; 2002US-0410957P.  
PR 17-SEP-2002; 2002US-0410958P.  
PR 17-SEP-2002; 2002US-0410959P.  
PR 17-SEP-2002; 2002US-0410960P.  
PR 17-SEP-2002; 2002US-0410961P.  
PR 17-SEP-2002; 2002US-0410962P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411023P.  
PR 17-SEP-2002; 2002US-0411024P.  
PR 17-SEP-2002; 2002US-0411032P.  
PR 17-SEP-2002; 2002US-0411035P.  
PR 17-SEP-2002; 2002US-0411037P.  
PR 17-SEP-2002; 2002US-0411041P.  
PR 17-SEP-2002; 2002US-0411045P.  
PR 17-SEP-2002; 2002US-0411046P.  
PR 17-SEP-2002; 2002US-0411048P.  
PR 17-SEP-2002; 2002US-0411052P.



PR 17-SEP-2002; 2002US-0411055P.  
PR 17-SEP-2002; 2002US-0411073P.  
PR 17-SEP-2002; 2002US-0411082P.  
PR 17-SEP-2002; 2002US-0411101P.  
PR 17-SEP-2002; 2002US-0411111P.  
PR 18-APR-2003; 2003US-0463700P.  
PR 18-APR-2003; 2003US-0463708P.  
PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471366P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493373P.  
PR 08-AUG-2003; 2003US-0493577P.

## (FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halembeck RF, Huang MM, Kotnakota S, Haislan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
PT genetic, bacterial and viral diseases.

## Claim 1: SEQ ID NO 2665; 428bp; English.

XX The present invention relates to an isolated nucleic acid molecule  
CC encoding a polypeptide which is believed to be cytostatic,  
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
CC composition and methods are useful for diagnosing, preventing and  
CC treating diseases such as proliferative (e.g. cancer), inflammatory,  
CC immune, metabolic, genetic, bacterial and viral diseases. The present  
CC sequence represents a human secreted protein. The present sequence is  
CC available on WIPOMB and is not in the specification.

XX Sequence 3411 AA;

Query Match 64.3%; Score 18; DB 8; Length 3411;  
Best Local Similarity 16.7%; Pred. No. 0.0001;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
Db 2826 CAATTATTTC 2837

## RESULT 47

ABU88256  
ID ABU88256 standard; protein; 4440 AA.

XX ABU88256;

DT 07-JUL-2003 (first entry)

XX Novel human secreted and transmembrane PRO polypeptide #4.  
DE Human; secreted and transmembrane protein: PRO; gene therapy;  
XX tumour necrosis factor-alpha release; TNF-alpha release;  
XX chondrocyte proliferation; chondrocyte differentiation; tumour;  
XX adrenal tumour; lung tumour; colon tumour; breast tumour;  
XX prostate tumour; rectal tumour; cervical tumour; liver tumour.  
OS Homo sapiens.  
XX US2003032127-A1.  
XX 13-FEB-2003.  
PD 26-JUN-2002; 2002US-00183012.  
XX 18-SEP-1997; 97US-0059263P.  
XX 18-SEP-1997; 97US-0059266P.  
XX 17-OCT-1997; 97US-0062250P.  
XX 21-OCT-1997; 97US-0063486P.  
XX 24-OCT-1997; 97US-0063120P.  
XX 24-OCT-1997; 97US-0063121P.  
XX 28-OCT-1997; 97US-0063540P.  
XX 28-OCT-1997; 97US-0063541P.  
XX 28-OCT-1997; 97US-0063544P.  
XX 28-OCT-1997; 97US-0063546P.  
XX 29-OCT-1997; 97US-0063734P.  
XX 31-OCT-1997; 97US-0063870P.  
XX 31-OCT-1997; 97US-0064103P.  
XX 13-NOV-1997; 97US-0065311P.  
XX 13-NOV-1997; 97US-0066120P.  
XX 24-NOV-1997; 97US-0066466P.  
XX 24-NOV-1997; 97US-0066772P.  
XX 11-DEC-1997; 97US-0069335P.  
XX 12-DEC-1997; 97US-0069425P.  
XX 17-DEC-1997; 97US-0069870P.  
XX 18-DEC-1997; 97US-0068017P.  
XX 10-MAR-1998; 98US-0077450P.  
XX 11-MAR-1998; 98US-0077632P.  
XX 11-MAR-1998; 98US-0077649P.  
XX 20-MAR-1998; 98US-0078886P.  
XX 20-MAR-1998; 98US-0078939P.  
XX 27-MAR-1998; 98US-0079664P.  
XX 27-MAR-1998; 98US-0079786P.  
XX 31-MAR-1998; 98US-0080107P.  
XX 31-MAR-1998; 98US-0080194P.  
XX 01-APR-1998; 98US-0080327P.  
XX 01-APR-1998; 98US-0080333P.  
XX 08-APR-1998; 98US-0081049P.  
XX 08-APR-1998; 98US-0081070P.  
XX 09-APR-1998; 98US-0081195P.  
XX 15-APR-1998; 98US-0081838P.  
XX 21-APR-1998; 98US-0082568P.  
XX 21-APR-1998; 98US-0082569P.  
XX 22-APR-1998; 98US-0082704P.  
XX 22-APR-1998; 98US-0082797P.  
XX 28-APR-1998; 98US-0083322P.  
XX 29-APR-1998; 98US-0083495P.  
XX 29-APR-1998; 98US-0083496P.  
XX 29-APR-1998; 98US-0083499P.  
XX 29-APR-1998; 98US-0083559P.  
XX 05-MAY-1998; 98US-0083566P.  
XX 06-MAY-1998; 98US-0084414P.  
XX 07-MAY-1998; 98US-0084639P.  
XX 07-MAY-1998; 98US-0084640P.  
XX 07-MAY-1998; 98US-0084643P.  
XX 15-MAY-1998; 98US-0085579P.  
XX 15-MAY-1998; 98US-0085580P.  
XX 15-MAY-1998; 98US-0085582P.  
XX 15-MAY-1998; 98US-0085700P.  
XX 18-MAY-1998; 98US-0086023P.  
XX 22-MAY-1998; 98US-0086392P.

PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088302P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 09-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088611P.  
PR 10-JUN-1998; 98US-0088624P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089388P.  
PR 17-JUN-1998; 98US-0089398P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 04-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095988P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-009691P.  
PR 17-AUG-1998; 98US-009697P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.

PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098832P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101923P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102310P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102587P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.

Query Match 64.3%; Score 18; DB 6; Length 4440;  
Best Local Similarity 16.7%; Pred. No. 0.00011;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXC 12  
Db 3858 CATTTTAATAC 3869

RESULT 48  
ABU90135 standard; protein; 4440 AA.  
ID ABU90135  
XX ABU90135;  
DT 11-AUG-2003 (first entry)  
XX  
DE Novel human secreted and transmembrane PRO protein #4.  
XX  
XX Human; gene therapy; tissue typing; tumour; chondrocyte proliferation;  
KW chondrocyte differentiation; tumour necrosis factor-alpha release;  
KW affinity purification.  
XX

OS Homo sapiens.  
XX  
PN US2003036147-A1.  
XX  
PD 20-FEB-2003.  
XX  
PF 02-JUL-2002; 2002US-00187741.  
XX  
PR 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-008579P.  
PR 15-MAY-1998; 98US-008580P.  
PR 15-MAY-1998; 98US-0085882P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.

PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 24-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095988P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096575P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-009687P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.

PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0098602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98US-0100849P.  
PR 17-SEP-1998; 98US-0100830P.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102230P.  
PR 29-SEP-1998; 98US-0102311P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.  
PR 07-OCT-1998; 98US-00168978.  
Query Match 64.3%; Score 18; DB 6; Length 4440;  
Best Local Similarity 16.7%; Pred. No. 0.00011;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 CXXXXXXXXXC 12  
DB 3858 CATTTTTATATAC 3869  
RESULT 49  
ABU96437  
ID ABU96437 strand: protein; 4440 AA.  
XX ABU96437;  
XX 25-JUL-2003 (first entry)  
XX  
XX Novel human secreted and transmembrane protein #4.  
XX  
XX Human; secreted and transmembrane protein; PRO; transgenic animal;  
XX knock; chromosome identification; tissue typing; tumour;  
XX chondrocyte proliferation; chondrocyte differentiation;  
XX tumor necrosis factor-alpha release stimulator.  
XX  
XX Homo sapiens.  
XX  
XX US2003036144-A1.  
XX  
XX 20-FEB-2003.  
XX

PF 01-JUL-2002; 2002US-00187601.  
XX  
XX 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 20-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080331P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081070P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 05-MAY-1998; 98US-0083559P.  
PR 06-MAY-1998; 98US-0084366P.  
PR 07-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 07-MAY-1998; 98US-0084643P.  
PR 15-MAY-1998; 98US-0085579P.  
PR 15-MAY-1998; 98US-0085580P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.

PR 09-JUN-1998; 98US-008655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089090P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 04-AUG-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 26-AUG-1998; 98US-0097522P.  
PR 26-AUG-1998; 98US-0097552P.  
PR 26-AUG-1998; 98US-0097554P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.

PR 10-SEP-1998; 98US-0099763P.  
PR 10-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.  
PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.  
PR 02-OCT-1998; 98US-0102965P.  
PR 06-OCT-1998; 98US-0103258P.  
PR 06-OCT-1998; 98US-0103449P.

Query Match 64.3%; Score 18; DB 6; Length 4440;  
Best Local Similarity 16.7%; Pred. No. 0.00011;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12

Db 3858 CATTTTAATAC 3869

RESULT 50

ABU99046  
ID ABU99046 standard; protein; 4440 AA.

XX  
AC ABU99046;

XX  
DT 01-ANG-2003 (first entry)

DE Novel human secreted and transmembrane protein #4.

XX Human; secreted and transmembrane protein; PRO; cytosolic; gene therapy;  
KW chondrocyte stimulator; tumour; adrenal tumour; lung tumour;  
KW colon tumour; breast tumour; prostate tumour; rectal tumour;  
KW cervical tumour; liver tumour; TNF-alpha release;  
KW tumour necrosis factor alpha release; chondrocyte cell proliferation;  
KW chondrocyte cell differentiation; pharmaceutical; diagnostic; biosensor;  
KW bioreactor.

XX Homo sapiens.

XX OS  
XX PN US2003013153-A1.

XX  
XX PD 16-JAN-2003.

XX PF 19-JUN-2002; 2002US-00175737.

XX  
XX 18-SEP-1997; 97US-0059263P.  
PR 18-SEP-1997; 97US-0059266P.

PR 17-OCT-1997; 97US-0062350P  
PR 21-OCT-1997; 97US-0063486P.  
PR 24-OCT-1997; 97US-0063120P.  
PR 24-OCT-1997; 97US-0063121P.  
PR 28-OCT-1997; 97US-0063540P.  
PR 28-OCT-1997; 97US-0063541P.  
PR 28-OCT-1997; 97US-0063544P.  
PR 28-OCT-1997; 97US-0063564P.  
PR 29-OCT-1997; 97US-0063734P.  
PR 31-OCT-1997; 97US-0063870P.  
PR 31-OCT-1997; 97US-0064103P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 21-NOV-1997; 97US-0066120P.  
PR 24-NOV-1997; 97US-0066466P.  
PR 24-NOV-1997; 97US-0066772P.  
PR 11-DEC-1997; 97US-0069335P.  
PR 12-DEC-1997; 97US-0069425P.  
PR 17-DEC-1997; 97US-0069870P.  
PR 18-DEC-1997; 97US-0068017P.  
PR 10-MAR-1998; 98US-0077450P.  
PR 11-MAR-1998; 98US-0077632P.  
PR 11-MAR-1998; 98US-0077649P.  
PR 20-MAR-1998; 98US-0078886P.  
PR 20-MAR-1998; 98US-0078939P.  
PR 27-MAR-1998; 98US-0079664P.  
PR 27-MAR-1998; 98US-0079786P.  
PR 31-MAR-1998; 98US-0080107P.  
PR 31-MAR-1998; 98US-0080194P.  
PR 01-APR-1998; 98US-0080327P.  
PR 01-APR-1998; 98US-0080333P.  
PR 08-APR-1998; 98US-0081049P.  
PR 08-APR-1998; 98US-0081070P.  
PR 09-APR-1998; 98US-0081195P.  
PR 15-APR-1998; 98US-0081838P.  
PR 21-APR-1998; 98US-0082568P.  
PR 21-APR-1998; 98US-0082569P.  
PR 22-APR-1998; 98US-0082704P.  
PR 22-APR-1998; 98US-0082797P.  
PR 28-APR-1998; 98US-0083322P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083495P.  
PR 29-APR-1998; 98US-0083496P.  
PR 29-APR-1998; 98US-0083499P.  
PR 29-APR-1998; 98US-0083559P.  
PR 05-MAY-1998; 98US-0084366P.  
PR 06-MAY-1998; 98US-0084414P.  
PR 07-MAY-1998; 98US-0084639P.  
PR 07-MAY-1998; 98US-0084640P.  
PR 15-MAY-1998; 98US-0085582P.  
PR 15-MAY-1998; 98US-0085700P.  
PR 18-MAY-1998; 98US-0086023P.  
PR 22-MAY-1998; 98US-0086392P.  
PR 22-MAY-1998; 98US-0086486P.  
PR 28-MAY-1998; 98US-0087098P.  
PR 28-MAY-1998; 98US-0087208P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088302P.  
PR 05-JUN-1998; 98US-0088322P.  
PR 05-JUN-1998; 98US-0088322P.  
PR 05-JUN-1998; 98US-0088322P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088722P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088740P.  
  
PR 10-JUN-1998; 98US-0088811P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088825P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088863P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0088909P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-00105413.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091010P.  
PR 01-JUL-1998; 98US-0091359P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091473P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091632P.  
PR 02-JUL-1998; 98US-0094006P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 10-AUG-1998; 98US-0095998P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 01-SEP-1998; 98US-0098716P.  
PR 01-SEP-1998; 98US-0098723P.  
PR 02-SEP-1998; 98US-0098803P.  
PR 02-SEP-1998; 98US-0098821P.  
PR 02-SEP-1998; 98US-0098843P.  
PR 09-SEP-1998; 98US-0099602P.  
PR 10-SEP-1998; 98US-0099741P.  
PR 10-SEP-1998; 98US-0099754P.  
PR 10-SEP-1998; 98US-0099763P.  
PR 15-SEP-1998; 98US-0099812P.  
PR 15-SEP-1998; 98US-0100388P.  
PR 16-SEP-1998; 98US-0100662P.

PR 16-SEP-1998; 98US-0100664P.  
PR 16-SEP-1998; 98US-0101751P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100683P.  
PR 17-SEP-1998; 98US-0100684P.  
PR 17-SEP-1998; 98US-0100919P.  
PR 17-SEP-1998; 98US-0100930P.  
PR 18-SEP-1998; 98US-0100849P.  
PR 18-SEP-1998; 98US-0101014P.  
PR 18-SEP-1998; 98US-0101068P.  
PR 23-SEP-1998; 98US-0101471P.  
PR 23-SEP-1998; 98US-0101472P.  
PR 23-SEP-1998; 98US-0101475P.  
PR 23-SEP-1998; 98US-0101477P.  
PR 24-SEP-1998; 98US-0101738P.  
PR 24-SEP-1998; 98US-0101739P.  
PR 24-SEP-1998; 98US-0101743P.  
PR 24-SEP-1998; 98US-0101922P.  
PR 25-SEP-1998; 98US-0101786P.  
PR 29-SEP-1998; 98US-0102207P.  
PR 29-SEP-1998; 98US-0102240P.  
PR 29-SEP-1998; 98US-0102330P.  
PR 29-SEP-1998; 98US-0102331P.  
PR 30-SEP-1998; 98US-0102487P.  
PR 30-SEP-1998; 98US-0102570P.  
PR 30-SEP-1998; 98US-0102571P.  
PR 01-OCT-1998; 98US-0102684P.  
PR 01-OCT-1998; 98US-0102687P.

Query Match 64.3%; Score 18; DB 6; Length 4440;  
Best Local Similarity 16.7%; Pred. No. 0.00011;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 3858 CATTTTAATAC 3869

Search completed: January 4, 2006, 15:56:18  
Job time : 83.1217 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:13:29 ; Search time 12.4174 Seconds  
(without alignments)  
92.983 Million cell updates/sec

Title: US-09-932-322-12  
Perfect score: 28  
Sequence: 1 CXXXXXXXXXC 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	64.3	307	1	GSFPR3
2	17	60.7	245	2	T24565
3	16	57.1	152	2	D72630
4	16	57.1	244	2	T31838
5	16	57.1	248	2	T31841
6	16	57.1	257	2	T21029
7	16	57.1	539	2	T52565
8	16	57.1	667	2	A48579
9	16	57.1	770	2	T51024
10	16	57.1	1041	2	T31097
11	15	53.6	154	2	AC0416
12	15	53.6	226	2	B41378
13	15	53.6	248	2	T21786
14	15	53.6	268	2	A84746
15	15	53.6	273	2	S69193
16	15	53.6	283	2	AH2744
17	15	53.6	287	2	S57770
18	15	53.6	289	2	S49812
19	15	53.6	318	2	G97525
20	15	53.6	357	2	T21152
21	15	53.6	363	2	S18765
22	15	53.6	392	2	T27303
23	15	53.6	409	2	T08928
24	15	53.6	420	2	S74388
25	15	53.6	423	2	D96552
26	15	53.6	423	2	S31126
27	15	53.6	569	2	B87180
28	15	53.6	727	2	UC7818
29	15	53.6	779	1	S40382

30	15	53.6	1003	1	PYVZAM
31	15	53.6	1043	2	T19734
32	15	53.6	1186	2	T18210
33	15	53.6	1364	2	T10236
34	15	53.6	1451	2	S65571
35	15	53.6	2098	2	T25888
36	15	53.6	2176	2	T13806
37	15	53.6	2187	2	S60224
38	15	53.6	2533	2	T28675
39	15	53.6	2533	2	T28674
40	15	53.6	2543	2	T31687
41	15	53.6	2703	1	A24420
42	15	53.6	4767	2	T31345
43	15	53.6	7829	2	T15789
44	15	53.6	7962	2	T18346
45	15	50.0	21	2	D22565
46	14	50.0	88	2	G82754
47	14	50.0	92	2	T54781
48	14	50.0	97	2	T28947
49	14	50.0	100	2	C72775
50	14	50.0	101	2	T14780
51	14	50.0	103	2	S73440
52	14	50.0	107	2	S46078
53	14	50.0	107	2	S57388
54	14	50.0	107	4	S67638
55	14	50.0	129	2	S48641
56	14	50.0	130	2	S57896
57	14	50.0	131	2	A43980
58	14	50.0	133	2	S44539
59	14	50.0	135	2	T15610
60	14	50.0	143	2	A54793
61	14	50.0	154	2	E72513
62	14	50.0	156	2	T31839
63	14	50.0	156	2	AH0865
64	14	50.0	156	2	C65065
65	14	50.0	158	2	S35201
66	14	50.0	159	2	E82692
67	14	50.0	164	2	T28741
68	14	50.0	165	2	T28742
69	14	50.0	166	2	T33970
70	14	50.0	177	2	S73288
71	14	50.0	177	2	S30932
72	14	50.0	177	2	F72706
73	14	50.0	195	2	S39967
74	14	50.0	197	2	T28739
75	14	50.0	199	2	S50363
76	14	50.0	204	2	T41853
77	14	50.0	205	2	T10384
78	14	50.0	209	2	UC7239
79	14	50.0	214	2	C70812
80	14	50.0	225	2	T22417
81	14	50.0	232	2	A60095
82	14	50.0	249	2	T31837
83	14	50.0	250	2	T31836
84	14	50.0	250	2	T31835
85	14	50.0	259	2	H71353
86	14	50.0	264	2	C70603
87	14	50.0	264	2	S42793
88	14	50.0	268	2	T05971
89	14	50.0	270	2	T27786
90	14	50.0	274	2	D82784
91	14	50.0	275	2	T05969
92	14	50.0	276	2	AE0860
93	14	50.0	286	1	C42053
94	14	50.0	288	2	T21158
95	14	50.0	293	2	T31840
96	14	50.0	298	2	S15524
97	14	50.0	300	2	T05729
98	14	50.0	300	2	S16344
99	14	50.0	312	2	T31834
100	14	50.0	315	2	T34526
101	14	50.0	328	2	T16065
102	14	50.0	329	2	T15823

spheroidin precurs  
hypothetical prote  
delta endotoxin -  
xanthine dehydroge  
pattern formation  
hypothetical prote  
toucan gene protei  
polyketide synthas  
alpha-51D immobili  
alpha-51D-immobil  
surface antigen - P  
notch protein - fr  
hypothetical prote  
hypothetical prote  
elastic titlin - hu  
R-phycocerythrin be  
hypothetical prote  
fibroblast growth  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
probable membrane  
hypothetical prote  
probable membrane  
lysosome (BC 3.2.1  
polyketide synthas  
neurophysin 2 [val  
probable membrane  
hypothetical prote  
pointiculin precurs  
hypothetical prote  
hypothetical prote  
preillin peptidase  
preillin peptidase  
serine proteinase  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
phycoerythrin beta  
R-phycocerythrin be  
hypothetical prote  
hypothetical prote  
hypothetical prote  
nitrlase homolog  
AckMPV orf115 - Bo  
peroxiredoxin V -  
probable lipQ prot  
hypothetical prote  
larval glue protei  
hypothetical prote  
hypothetical prote  
hypothetical prote  
conserved hypothet  
hypothetical prote  
PES-1 protein - Ca  
probable receptor-  
phosphomethylpyrim  
receptor-like kina  
hypothetical prote  
gap junction prote  
hypothetical prote  
hypothetical prote  
variant surface gl  
hypothetical prote  
hypothetical prote  
iron-sulfur cofact  
hypothetical prote



249	13	46.4	66	2	A28644	kappa-neurotoxin -	322	13	46.4	201	2	T16181	hypothetical prote
250	13	46.4	66	2	JN0408	neurotoxin Os-1 -	323	13	46.4	202	2	C90742	hypothetical prote
251	13	46.4	71	2	T00008	copy number contro	324	13	46.4	202	2	F85592	hypothetical prote
252	13	46.4	79	2	T27605	hypothetical prote	325	13	46.4	204	2	D72864	AcOf1-115 protein
253	13	46.4	81	2	B42465	probable transcrip	326	13	46.4	205	2	H71457	hypothetical prote
254	13	46.4	98	2	E84599	hypothetical prote	327	13	46.4	208	2	C97654	hypothetical prote
255	13	46.4	101	2	T42462	hypothetical prote	328	13	46.4	208	2	A12877	conserved hypochet
256	13	46.4	102	2	AE1357	hypothetical prote	329	13	46.4	209	2	A82680	2-amino-4-hydroxy-
257	13	46.4	102	2	AE1727	hypothetical prote	330	13	46.4	210	2	S76973	hypothetical prote
258	13	46.4	107	2	T27988	hypothetical prote	331	13	46.4	213	2	F75585	urase accessory p
259	13	46.4	108	2	B95314	hypothetical prote	332	13	46.4	216	2	T50615	hypothetical prote
260	13	46.4	111	2	AC0733	probable exported	333	13	46.4	216	2	C90162	acetyltransferase,
261	13	46.4	111	2	T20452	hypothetical prote	334	13	46.4	217	2	A86600	CT734 hypothetical
262	13	46.4	111	2	T34565	hypothetical prote	335	13	46.4	217	2	A72024	conserved hypochet
263	13	46.4	112	2	T51142	nonspecific lipid-	336	13	46.4	217	2	A42644	neut protein - Esc
264	13	46.4	112	2	T48770	hypothetical prote	337	13	46.4	217	2	T48642	hypothetical prote
265	13	46.4	114	2	S57692	probable membrane	338	13	46.4	217	2	UC7875	l-rhamnose-binding
266	13	46.4	114	2	F69264	hypothetical prote	339	13	46.4	220	2	T25789	hypothetical prote
267	13	46.4	114	2	T01930	ethylene-response	340	13	46.4	222	2	T27939	hypothetical prote
268	13	46.4	117	2	T02169	homeobox protein H	341	13	46.4	225	2	S45356	probable serine pr
269	13	46.4	119	2	B45937	early chorion pro-	342	13	46.4	226	2	G95287	hypochet (imported
270	13	46.4	119	2	S24292	chorion protein -	343	13	46.4	226	2	T21380	hypothetical prote
271	13	46.4	119	2	S24294	chorion protein -	344	13	46.4	227	2	A33937	Ig light chain (13
272	13	46.4	119	2	S24291	chorion protein -	345	13	46.4	227	2	A55472	ornithine decarbox
273	13	46.4	121	2	S24293	mercury resistance	346	13	46.4	229	1	MMVYPM	25k protein - puta
274	13	46.4	125	2	S30170	probable resistance	347	13	46.4	233	2	T21496	hypothetical prote
275	13	46.4	126	2	AB0071	nucleic acid-bind	348	13	46.4	235	2	T16916	hypothetical prote
276	13	46.4	127	2	S27915	hypothetical prote	349	13	46.4	236	2	E84307	glucose 1-dehydrog
277	13	46.4	127	2	H90091	hypothetical prote	350	13	46.4	237	2	T25152	hypothetical prote
278	13	46.4	129	2	T47958	hypothetical prote	351	13	46.4	238	2	S77699	inner cell wall ma
279	13	46.4	132	1	MFIV2J	matrix protein M2	352	13	46.4	238	2	T22098	hypothetical prote
280	13	46.4	141	2	T33983	hypothetical prote	353	13	46.4	241	2	T24782	hypothetical prote
281	13	46.4	144	2	A36324	growth arrest-spec	354	13	46.4	242	2	A96794	unknown protein fl
282	13	46.4	145	2	H87436	hypothetical prote	355	13	46.4	243	2	H65136	hypothetical prote
283	13	46.4	145	2	T33037	hypothetical prote	356	13	46.4	243	2	F86006	hypothetical prote
284	13	46.4	146	1	W9BPT3	gene 19.2 protein	357	13	46.4	243	2	G91160	hypothetical prote
285	13	46.4	148	2	S19462	probable membrane	358	13	46.4	246	2	T00628	hypothetical prote
286	13	46.4	150	2	S66926	hypothetical prote	359	13	46.4	246	2	T25532	hypothetical prote
287	13	46.4	150	2	C82890	conserved hypochet	360	13	46.4	247	2	T21406	hypothetical prote
288	13	46.4	156	2	C91089	prepillin peptidase	361	13	46.4	247	2	S72859	hypothetical prote
289	13	46.4	156	2	D75258	hypothetical prote	362	13	46.4	248	2	S05007	calcium channel al
290	13	46.4	156	2	E85934	prepillin peptidase	363	13	46.4	248	2	T19913	hypothetical prote
291	13	46.4	157	2	JN0057	hypothetical 17k p	364	13	46.4	249	2	T21785	hypothetical prote
292	13	46.4	160	2	JN0503	peripheral myelin	365	13	46.4	249	2	T27782	hypothetical prote
293	13	46.4	160	2	S21721	growth arrest-spec	366	13	46.4	250	2	T36131	hypothetical prote
294	13	46.4	160	2	A41144	growth arrest-reja	367	13	46.4	250	2	T28725	hypothetical prote
295	13	46.4	162	2	T11978	phycoerythrin alpha	368	13	46.4	252	2	T28527	hypothetical prote
296	13	46.4	163	2	D83883	hypothetical prote	369	13	46.4	254	2	E83619	hypochet
297	13	46.4	163	2	B44813	probable molybdenu	370	13	46.4	255	2	T03408	MADS box protein -
298	13	46.4	166	1	NVB02	vasopressin / neur	371	13	46.4	255	2	T34163	hypothetical prote
299	13	46.4	166	2	F81696	hypothetical prote	372	13	46.4	257	2	D71544	hypothetical prote
300	13	46.4	166	2	T32701	hypothetical prote	373	13	46.4	259	2	T27783	hypothetical prote
301	13	46.4	168	1	NVR72	vasopressin / neur	374	13	46.4	261	2	S70006	finger protein zfo
302	13	46.4	168	2	B43755	vasopressin / neur	375	13	46.4	262	2	S01914	El membrane glycop
303	13	46.4	168	2	T38944	melanoma ubiquitin	376	13	46.4	262	2	G86938	conserved hypochet
304	13	46.4	171	2	H84923	hypothetical prote	377	13	46.4	265	2	T30059	hypochet
305	13	46.4	173	2	P96543	hypothetical prote	378	13	46.4	266	2	T10816	thyroxine deiodina
306	13	46.4	173	2	A71144	hypothetical prote	379	13	46.4	268	2	T43486	hypothetical prote
307	13	46.4	176	2	T30961	hypothetical prote	380	13	46.4	269	2	T21407	hypothetical prote
308	13	46.4	177	2	A95283	hypochet	381	13	46.4	271	2	E86352	protein T26F17.16
309	13	46.4	180	2	PS0011	lignin peroxidase	382	13	46.4	272	2	T19418	hypochet
310	13	46.4	180	2	E71337	hypochet	383	13	46.4	274	2	T39087	T-cell surface pro
311	13	46.4	182	2	S29611	pollen-preferentia	384	13	46.4	275	2	A34866	hypochet
312	13	46.4	183	2	T38923	microfibril-associ	385	13	46.4	275	2	S08464	T-cell alloantigen
313	13	46.4	183	2	A54151	microfibril-associ	386	13	46.4	275	2	T05822	hypochet
314	13	46.4	183	2	A49313	microfibril-associ	387	13	46.4	276	2	T43541	purine nucleotide
315	13	46.4	185	2	S57419	PAR-1a protein - c	388	13	46.4	276	2	AD1987	hypochet
316	13	46.4	185	2	S62698	phocassimilate-re	389	13	46.4	277	2	H97266	mind family ATPase
317	13	46.4	185	2	S62699	phocassimilate-re	390	13	46.4	278	2	T16494	hypochet
318	13	46.4	198	2	UB0316	procozoan RNase, Nd	391	13	46.4	280	2	A42424	chitinase (EC 3.2.
319	13	46.4	199	2	T48099	hypochet	392	13	46.4	287	1	A45168	probable 3-oxoacyl
320	13	46.4	199	2	T32267	hypochet	393	13	46.4	287	2	T04236	xyloglucan endo-1,
321	13	46.4	201	2	S37847	hypochet	394	13	46.4	289	2	AE2006	hypochet

395	13	46.4	291	2	T28846	hypotheical prote	468	13	46.4	380	1	D65201	glycerol dehydroge
396	13	46.4	294	1	G65106	lipoprotein precu	469	13	46.4	380	2	B91238	glycerol dehydroge
397	13	46.4	294	2	AC0901	conserved hypothe	470	13	46.4	380	2	B66085	glycerol dehydroge
398	13	46.4	294	2	D91134	lipoprotein precu	471	13	46.4	382	2	B86268	FL13B4.1 protein -
399	13	46.4	294	2	G85979	lipoprotein precu	472	13	46.4	384	2	S25771	gaal protein - mou
400	13	46.4	294	2	T23916	hypotheical prote	473	13	46.4	389	2	AD0722	glutamate dehydrog
401	13	46.4	294	2	T23682	hypotheical prote	474	13	46.4	393	2	A56947	dual specificity p
402	13	46.4	296	2	UT0548	Killer toxin KHR p	475	13	46.4	394	2	A56115	dual specificity p
403	13	46.4	297	2	S06267	surface antigen H	476	13	46.4	395	2	AB0814	probable alcohol
404	13	46.4	300	2	T48147	hypotheical prote	477	13	46.4	396	2	B49425	Desert hedgehog pr
405	13	46.4	301	2	T05493	pathogenesis-relat	478	13	46.4	400	2	B48613	surface protein su
406	13	46.4	301	2	T40593	cytoplasmic dynein	479	13	46.4	401	2	B87684	hypotheical prote
407	13	46.4	303	2	H71277	probable DNA adent	480	13	46.4	403	2	A13256	hypotheical prote
408	13	46.4	303	2	T46405	hypotheical prote	481	13	46.4	404	2	S58830	probable membrane
409	13	46.4	303	2	B88990	protein C36C5.13 l	482	13	46.4	410	2	AB3546	aminobutyraldehyde
410	13	46.4	309	2	S37934	hypotheical prote	483	13	46.4	411	2	AB2760	cell division prot
411	13	46.4	311	2	H64483	5-methyltetrahydro	484	13	46.4	412	2	T17653	hypotheical prote
412	13	46.4	312	2	B75405	endopeptidase-rela	485	13	46.4	412	2	C38104	LFY floral meriste
413	13	46.4	313	2	AC0603	asparaginase (EC 3	486	13	46.4	415	2	C84829	hypotheical prote
414	13	46.4	314	1	AS7126	dual specificity p	487	13	46.4	415	2	T43352	nuclear receptor N
415	13	46.4	314	1	B57126	regenerating liver	488	13	46.4	416	2	S65110	chitinase (EC 3.2.
416	13	46.4	314	2	A44437	probable elicitor	489	13	46.4	420	2	B38104	LFY floral meriste
417	13	46.4	317	2	T00500	protein F7H2.14 i1	490	13	46.4	422	2	S35197	hypotheical prote
418	13	46.4	317	2	D86292	probable asparagin	491	13	46.4	424	2	A38104	LFY floral meriste
419	13	46.4	321	2	D64820	cyserine proteinas	492	13	46.4	425	2	B84631	probable serine ca
420	13	46.4	323	2	S19650	autocrine motility	493	13	46.4	425	2	F70608	hypotheical prote
421	13	46.4	323	2	A39677	alkanal monooxygen	494	13	46.4	426	2	F85574	probable symport p
422	13	46.4	324	1	D42951	cathepsin L (EC 3.	495	13	46.4	426	2	F90723	probable transport
423	13	46.4	324	2	S47432	hypotheical prote	496	13	46.4	430	2	JC4720	toxin co-regulated
424	13	46.4	324	2	C86582	conserved hypothe	497	13	46.4	433	1	S24353	proteasome 26S sub
425	13	46.4	324	2	D72041	conserved hypothe	498	13	46.4	433	2	S53709	MS1 protein homol
426	13	46.4	325	2	T27688	hypotheical prote	499	13	46.4	433	2	S39349	hypocetical prote
427	13	46.4	326	4	S61652	hypotheical prote	500	13	46.4	433	2	F86163	hypocetical prote
428	13	46.4	327	2	T41463	conserved hypothe	501	13	46.4	434	2	S50865	avemectin-sensiti
429	13	46.4	329	2	S74227	cathepsin K (EC 3.	502	13	46.4	435	2	T20152	hypotheical prote
430	13	46.4	330	2	T25323	zinc-finger protei	503	13	46.4	436	2	B81033	N-acetylglutamate
431	13	46.4	335	2	JB0115	Viril1 homolog - B	504	13	46.4	436	2	AB1977	probable amino-ac
432	13	46.4	339	2	F47301	phosphoprotein pho	505	13	46.4	437	2	S11497	branched-chain ami
433	13	46.4	341	1	T22930	hypotheical prote	506	13	46.4	437	2	H82295	branched chain ami
434	13	46.4	341	1	C70611	hemagglutinin - in	507	13	46.4	438	2	I38946	melanoma ubiqlitou
435	13	46.4	344	1	JQ1643	N4- (beta-N-acetyl	508	13	46.4	440	2	D69596	branched-chain ami
436	13	46.4	346	1	MUHGCD	calcium channel pr	509	13	46.4	440	2	F97098	branched-chain ami
437	13	46.4	347	2	S12955	NADH2 dehydrogen	510	13	46.4	445	2	S37779	porin precursor, m
438	13	46.4	348	2	T14141	adenine glycosylas	511	13	46.4	445	2	C69596	hypocetical prote
439	13	46.4	349	2	B81855	A/G-specific adeni	512	13	46.4	445	2	T18129	allantoinase homol
440	13	46.4	349	2	D81088	hypotheical prote	513	13	46.4	446	1	S50546	adenosylhomocyste
441	13	46.4	353	2	T27800	NADH2 dehydrogen	514	13	46.4	449	2	T09745	myb-related protei
442	13	46.4	354	2	S70595	fructose-bisphosph	515	13	46.4	453	2	D75446	oxidoreductase - D
443	13	46.4	355	1	ADZM	CONSTANS-like 1 -	516	13	46.4	454	2	T48973	cytochrome P450-1i
444	13	46.4	355	2	TS1414	fructose bisphosph	517	13	46.4	455	2	T48973	isochorismate synt
445	13	46.4	358	2	T47550	probable fructose	518	13	46.4	455	2	AG0308	triacylglycerol 1i
446	13	46.4	358	2	A84781	manganese peroxida	519	13	46.4	457	2	I48206	cellobiohydrolase
447	13	46.4	361	2	S59531	probable pectin ac	520	13	46.4	459	2	JC7931	triacylglycerol 1i
448	13	46.4	362	2	T05824	C-Fringe protein 1	521	13	46.4	461	2	S21223	triacylglycerol 1i
449	13	46.4	363	2	JCS536	phosphoprotein pho	522	13	46.4	461	2	S50864	avemectin-sensiti
450	13	46.4	364	1	T18972	40.7 kDa polypepti	523	13	46.4	464	2	C46157	hexokinase (EC 2.7
451	13	46.4	364	1	C45574	protein F13A2.6 i1	524	13	46.4	465	2	A46157	hexokinase (EC 2.7
452	13	46.4	364	2	H86384	dual specificity p	525	13	46.4	465	2	I49693	hexokinase (EC 2.7
453	13	46.4	366	2	H89025	dual specificity p	526	13	46.4	465	2	A31810	hexokinase (EC 2.7
454	13	46.4	367	1	S24411	dual specificity p	527	13	46.4	466	2	S41084	triacylglycerol 1i
455	13	46.4	367	1	S29090	dual specificity p	528	13	46.4	466	2	B46157	hexokinase (EC 2.7
456	13	46.4	367	1	MOHU	alpha-2-HS-glycop	529	13	46.4	467	2	E89605	protein P18G5.2 i1
457	13	46.4	367	2	AH0936	glycerol dehydroge	530	13	46.4	471	2	S54460	hypotheical prote
458	13	46.4	367	2	S52265	dual specificity p	531	13	46.4	472	1	A26730	ovotininhibitor preu
459	13	46.4	367	2	D82763	UDP-N-acetylglucos	532	13	46.4	474	2	B46746	glycine hydroxymet
460	13	46.4	367	2	JC6087	helix-loop-helix t	533	13	46.4	474	2	AF0759	cobyrinic acid A,C
461	13	46.4	368	2	H71600	rifin PBB101ow - m	534	13	46.4	475	2	A33696	glycine hydroxymet
462	13	46.4	371	2	T07938	probable acetyl-Co	535	13	46.4	475	2	A71302	conserved hypothe
463	13	46.4	372	1	OPUGAP	lignin peroxidase	536	13	46.4	479	2	T49871	peroxisomal Ca-dep
464	13	46.4	372	2	B32322	lignin peroxidase	537	13	46.4	482	2	T48397	S-receptor kinase-
465	13	46.4	372	2	JH0156	brefeldin A estera	538	13	46.4	487	1	S71770	calcium-dependent
466	13	46.4	372	2	A53050	heat shock transcr	539	13	46.4	487	2	A43758	aromatic-L-amino-a
467	13	46.4	377	2	T04213		540	13	46.4	488	2	D70876	probable polyketid

541	13	46.4	488	2	T30602	hypothetical prote	614	13	46.4	650	2	S44806	F10B9.6 protein -
542	13	46.4	489	2	H69059	dihydroxy-acid deh	615	13	46.4	654	2	C87791	protein B0207.12 l
543	13	46.4	494	2	T09400	probable cobQ prot	616	13	46.4	656	2	D84359	thermosome subunit
544	13	46.4	495	2	T20754	hypothetical prote	617	13	46.4	656	2	E29826	hypothetical 72.4k
545	13	46.4	496	2	T33496	hypothetical prote	618	13	46.4	657	2	T52460	hypothetical prote
546	13	46.4	498	2	S12061	hexokinase (EC 2.7	619	13	46.4	659	2	A57532	Na+-dependent nucl
547	13	46.4	498	2	S27849	variant surface gl	620	13	46.4	663	2	T30621	hypothetical prote
548	13	46.4	504	2	T01603	hypothetical prote	621	13	46.4	664	2	S60062	hevin precursor -
549	13	46.4	504	2	E86343	T2211.12 protein	622	13	46.4	669	2	A46511	envelope protein -
550	13	46.4	504	2	T16526	hypothetical prote	623	13	46.4	676	2	G84663	hypothetical prote
551	13	46.4	506	2	S37583	RING finger protei	624	13	46.4	677	2	C42125	trophozoite cystei
552	13	46.4	506	2	T29968	hypothetical prote	625	13	46.4	684	2	G84730	Mutator-like trans
553	13	46.4	509	2	A95985	probable sugar kin	626	13	46.4	690	2	T11749	transferrin - Acta
554	13	46.4	511	2	A95985	cellulotagmin I sylv	627	13	46.4	691	2	T31552	hypothetical prote
555	13	46.4	513	1	TVHURF	ret finger protein	628	13	46.4	700	2	S38426	chaperonin 60 proc
556	13	46.4	513	1	T19962	hypothetical prote	629	13	46.4	713	2	A35502	major surface-labe
557	13	46.4	520	2	A96922	beta-glucosidase h	630	13	46.4	730	2	A75486	hypothetical prote
558	13	46.4	530	2	UC7979	cellobiohydrolase	631	13	46.4	733	1	A46373	probable serine/ch
559	13	46.4	536	2	T24000	hypothetical prote	632	13	46.4	738	2	D86345	hypothetical prote
560	13	46.4	538	2	S52472	cell fusion protei	633	13	46.4	746	2	S52770	subtilisin-like se
561	13	46.4	539	2	A23923	carboxylesterase (	634	13	46.4	757	2	UC7519	hypothetical prote
562	13	46.4	539	2	T46132	hypothetical prote	635	13	46.4	757	2	T16609	3-oxoacyl-(acyl-ca
563	13	46.4	540	2	A31584	carboxylesterase (	636	13	46.4	764	2	H83879	integrin beta-8 ch
564	13	46.4	541	2	T33583	hypothetical prote	637	13	46.4	768	2	B41029	integrin beta-8 ch
565	13	46.4	541	2	B82342	probable transcrip	638	13	46.4	773	2	JH0609	protein-tyrosine-p
566	13	46.4	543	2	T27000	hypothetical prote	639	13	46.4	775	2	T00962	hypothetical prote
567	13	46.4	546	2	T49931	hypothetical prote	640	13	46.4	781	2	S43534	integrin beta3 - c
568	13	46.4	549	2	JX0054	carboxylesterase (	641	13	46.4	795	2	T34468	hypothetical prote
569	13	46.4	554	2	A39060	carboxylesterase (	642	13	46.4	796	2	UC1285	protein-tyrosine-p
570	13	46.4	557	2	A47162	thiolesterase B (B	643	13	46.4	799	2	A38308	integrin beta-5 ch
571	13	46.4	557	2	A48434	variant-specific s	644	13	46.4	802	2	A36065	protein-tyrosine-p
572	13	46.4	560	1	F69059	arginine-cRNA liga	645	13	46.4	807	2	C85025	hypothetical prote
573	13	46.4	561	2	S71597	carboxylesterase (	646	13	46.4	814	2	T49207	receptor kinase-11
574	13	46.4	561	2	S62788	carboxylesterase (	647	13	46.4	829	1	A47373	protein-tyrosine-p
575	13	46.4	561	2	C64459	dihydroxy-acid deh	648	13	46.4	832	2	JC9051	protein tyrosine p
576	13	46.4	562	2	A55281	carboxylesterase (	649	13	46.4	840	2	S41218	HTR1 protein - yea
577	13	46.4	563	1	VCWMV7	env polypepten -	650	13	46.4	843	1	JQ2229	DNA-directed DNA p
578	13	46.4	564	2	S08199	cytochrome-c3 hydr	651	13	46.4	848	1	T02053	S-receptor kinase
579	13	46.4	565	2	S10367	carboxylesterase (	652	13	46.4	852	2	S28415	guanine nucleotide
580	13	46.4	566	2	S19307	carboxylesterase (	653	13	46.4	852	2	T28790	hypothetical prote
581	13	46.4	566	2	S69889	hemagglutinin prec	654	13	46.4	855	2	UC7898	3',5'-cyclic nucle
582	13	46.4	567	1	A41010	carboxylesterase (	655	13	46.4	881	2	T21942	hypothetical prote
583	13	46.4	567	2	T08405	hexose transport p	656	13	46.4	892	2	A41697	nitrate assimilat
584	13	46.4	569	2	S50771	hypothetical prote	657	13	46.4	895	2	T23191	hypothetical prote
585	13	46.4	571	1	S30253	GABA transport pro	658	13	46.4	907	2	T27317	hypothetical prote
586	13	46.4	573	2	T27671	hypothetical prote	659	13	46.4	917	1	S15885	hexokinase (EC 2.7
587	13	46.4	579	2	T30635	hypothetical prote	660	13	46.4	917	2	UC2025	hexokinase (EC 2.7
588	13	46.4	581	2	T24393	hypothetical prote	661	13	46.4	931	2	T25148	hypothetical prote
589	13	46.4	582	1	VCVDAR	env polypepten -	662	13	46.4	934	2	G86548	polymorphic outer
590	13	46.4	591	2	T48141	acroganin - guine	663	13	46.4	934	2	A72075	hypothetical prote
591	13	46.4	594	2	H48613	env polypepten pr	664	13	46.4	950	2	F86222	hypothetical prote
592	13	46.4	596	2	T30498	probable ribonucle	665	13	46.4	951	2	T00260	hypothetical prote
593	13	46.4	601	2	D89711	protein P40E10.4 (	666	13	46.4	952	2	DB1593	polymorphic membra
594	13	46.4	601	2	T22025	hypothetical prote	667	13	46.4	962	2	UC5571	subtilisin-like pr
595	13	46.4	602	1	S25316	hydroxymethylgluta	668	13	46.4	966	2	B84053	penicillin-binding
596	13	46.4	602	2	T37254	acetylcholinestera	669	13	46.4	969	1	A39490	subtilisin-like pr
597	13	46.4	605	2	S48940	hypothetical prote	670	13	46.4	975	2	JC5570	subtilisin-like pr
598	13	46.4	609	2	UC7819	metalloproteinase	671	13	46.4	983	2	T39902	translacion Elonga
599	13	46.4	616	2	T29234	hypothetical prote	672	13	46.4	993	2	S35633	DNA-binding protei
600	13	46.4	618	2	G82340	phosphoglucanate d	673	13	46.4	994	2	H96510	probable disease r
601	13	46.4	625	2	T10661	serine/chreonine-s	674	13	46.4	1011	2	T50344	poly(A)+ RNA trans
602	13	46.4	626	2	B97358	glucose-inhibited	675	13	46.4	1012	2	T52000	poly(A)+ RNA trans
603	13	46.4	632	2	S58152	hypothetical prote	676	13	46.4	1016	2	T00375	hypothetical prote
604	13	46.4	632	2	T21602	hypothetical prote	677	13	46.4	1046	2	A26838	prestalk protein p
605	13	46.4	633	2	T24898	hypothetical prote	678	13	46.4	1060	2	T30823	bunetanide sensiti
606	13	46.4	634	1	GERTX1	matrix glycoprotei	679	13	46.4	1073	2	E83051	cardanoylphosphate
607	13	46.4	635	2	S76371	hypothetical prote	680	13	46.4	1074	2	UC5928	semaphorin F precu
608	13	46.4	636	2	H87789	protein C34G6.1 (I	681	13	46.4	1082	2	T31112	ATPase 2 (EC 3.6.1
609	13	46.4	638	2	T26490	hypothetical prote	682	13	46.4	1087	2	S02035	period clock prote
610	13	46.4	640	1	HHKW7A	dnak-type molecula	683	13	46.4	1093	2	T50652	AP-3 complex beta3
611	13	46.4	640	2	T07923	aceetyl-CoA carboxy	684	13	46.4	1094	2	T50651	AP3-complex beta-3
612	13	46.4	640	2	AB2251	glucose inhibited	685	13	46.4	1101	2	T16840	hypothetical prote
613	13	46.4	640	2	S62747	homeotic protein A	686	13	46.4	1105	2	C87622	Tomb-dependent rec

687	13	46.4	1112	2	H97050	DNA Polymerase III	760	13	46.4	2143	2	JH0427	voltage-dependent
688	13	46.4	1129	2	T42732	A-kinase anchoring	761	13	46.4	2143	2	G96595	hypothetical prote
689	13	46.4	1149	2	I38006	M130 antigen precu	762	13	46.4	2166	2	S11339	calcium channel pr
690	13	46.4	1151	2	I38004	M130 antigen precu	763	13	46.4	2168	2	T30171	nilein - mouse
691	13	46.4	1156	2	I38005	M130 antigen precu	764	13	46.4	2171	2	S05054	calcium channel al
692	13	46.4	1159	2	I38465	probable potassium	765	13	46.4	2180	2	T29764	hypothetical prote
693	13	46.4	1160	2	T13713	betal protein - fr	766	13	46.4	2188	2	A70984	probable polyketid
694	13	46.4	1162	2	T21557	hypothetical prote	767	13	46.4	2195	2	T34264	hypothetical prote
695	13	46.4	1166	2	T13958	synCAP-B1 protein	768	13	46.4	2201	2	S73014	polyketide synthas
696	13	46.4	1170	2	I45914	integrin alpha 2 s	769	13	46.4	2204	1	RKNZNV	genome polyprotein
697	13	46.4	1176	2	T49482	hypothetical prote	770	13	46.4	2220	1	A45290	calcium channel pr
698	13	46.4	1180	2	A35854	integrin alpha-1 c	771	13	46.4	2233	2	T28669	surface protein 51
699	13	46.4	1181	2	A33998	integrin alpha-2 c	772	13	46.4	2240	2	T37057	probable mult-dom
700	13	46.4	1205	2	T41987	hypothetical prote	773	13	46.4	2257	2	D66483	protein f55.19 [l
701	13	46.4	1205	2	S14201	probable adenylate	774	13	46.4	2257	2	T09538	acetyl-CoA carboxy
702	13	46.4	1249	2	T14270	Ras-GRase activat	775	13	46.4	2261	2	T07084	acetyl-CoA carboxy
703	13	46.4	1265	2	S57968	Ran-binding protei	776	13	46.4	2304	2	T07920	probable acetyl-Co
704	13	46.4	1271	2	T24008	hypothetical prote	777	13	46.4	2359	2	E86483	probable acetyl-Co
705	13	46.4	1274	2	T10729	transferrin-like p	778	13	46.4	2518	2	A12140	polyketide synthas
706	13	46.4	1274	2	T42017	cysteine rich proc	779	13	46.4	2524	2	A35844	Xorch protein - Af
707	13	46.4	1293	2	T14259	ras GRase-activat	780	13	46.4	2531	2	S18188	notch protein homo
708	13	46.4	1299	2	T43251	furin (EC 3.4.21.7	781	13	46.4	2531	2	A46019	polyketide synthas
709	13	46.4	1323	2	PN0568	connecin 3B - chl	782	13	46.4	2543	2	F69679	notch protein homo
710	13	46.4	1323	2	E88257	protein let-23 [lm	783	13	46.4	2555	2	A40043	hypothetical prote
711	13	46.4	1324	2	S06187	RNA2 polypotein -	784	13	46.4	2610	2	T20968	variant-specific s
712	13	46.4	1331	2	S05011	calcium channel al	785	13	46.4	2706	2	T28155	hypothetical prote
713	13	46.4	1340	2	D96521	protein F21D18.16	786	13	46.4	2844	2	S28291	fibillin-2 precu
714	13	46.4	1359	2	T10235	xanthine dehydroge	787	13	46.4	2918	2	A54105	hypothetical prote
715	13	46.4	1369	2	T43433	alpha-glucan synh	788	13	46.4	2946	2	T15840	protein HMWPI - Ye
716	13	46.4	1374	2	S70712	protein-tyrosine k	789	13	46.4	3161	2	T30342	protein HMWPI - Ye
717	13	46.4	1391	2	T20406	hypothetical prote	790	13	46.4	3163	2	AB0233	Yersinabactin bio
718	13	46.4	1402	2	D70634	probable polyketid	791	13	46.4	3224	1	T17440	probable polyketid
719	13	46.4	1407	2	S59823	probable membrane	792	13	46.4	3224	1	S58884	Ran-binding protei
720	13	46.4	1416	2	E88550	protein ZC84.1 [lm	793	13	46.4	3461	2	S58870	reelin precursor
721	13	46.4	1446	2	S73013	polyketide synthas	794	13	46.4	3507	2	T34513	hypothetical prote
722	13	46.4	1461	2	T26327	hypothetical prote	795	13	46.4	3828	2	T13857	cytothox protein
723	13	46.4	1484	2	T42632	breast cancer tumo	796	13	46.4	3971	2	T44806	mycosubtilin synth
724	13	46.4	1488	2	AG2136	polyketide synthas	797	13	46.4	4162	2	T42633	connectin/tictin -
725	13	46.4	1488	2	C70984	probable ppsb prot	798	13	46.4	5376	2	T42215	zonadhesin - mouse
726	13	46.4	1489	2	S73015	polyketide synthas	799	13	46.4	6420	2	T30283	polyketide synthas
727	13	46.4	1538	2	E70874	probable ppsb prot	800	12	42.9	35	2	S55030	CAP5 protein - ant
728	13	46.4	1540	2	H87203	polyketide synthas	801	12	42.9	27	2	A59356	neurotoxin Bmk 41-
729	13	46.4	1565	2	AD2135	polyketide synthas	802	12	42.9	40	1	SMFP	metallothionein Mt
730	13	46.4	1567	2	T03730	antigen containing	803	12	42.9	40	2	B61194	metallothionein Mt
731	13	46.4	1570	2	AC2012	hypothetical prote	804	12	42.9	41	2	B44010	kit-ligand (altern
732	13	46.4	1587	2	AB2012	hypothetical prote	805	12	42.9	43	2	D84641	hypothetical prote
733	13	46.4	1597	1	BVPSFL	sol protein, large	806	12	42.9	45	1	XKPO2A	proteinase inhibit
734	13	46.4	1597	2	T08428	gene small optic 1	807	12	42.9	46	2	S50015	leech-derived cryp
735	13	46.4	1722	2	E89753	protein PliC7.4 [l	808	12	42.9	50	2	F82539	hypothetical prote
736	13	46.4	1743	2	T26859	hypothetical prote	809	12	42.9	52	2	T10299	conotoxin-like pro
737	13	46.4	1774	2	T17421	polyketide synthas	810	12	42.9	54	2	A31436	ovomucoid, thrid d
738	13	46.4	1795	2	D97312	hypothetical prote	811	12	42.9	54	2	I61589	ovomucoid (PSTI-cy
739	13	46.4	1815	2	S73021	polyketide synthas	812	12	42.9	54	2	A61589	ovomucoid (PSTI-cy
740	13	46.4	1822	2	F87203	polyketide synthas	813	12	42.9	54	2	I61588	ovomucoid (PSTI-cy
741	13	46.4	1827	2	B70984	probable polyketid	814	12	42.9	54	2	F31439	ovomucoid, thrid d
742	13	46.4	1844	1	RRWPTM	genome polyprotein	815	12	42.9	54	2	I31446	ovomucoid, thrid d
743	13	46.4	1844	1	S01956	hypothetical prote	816	12	42.9	54	2	C31447	ovomucoid, thrid d
744	13	46.4	1871	2	A87204	polyketide synthas	817	12	42.9	54	2	D61589	ovomucoid (PSTI-cy
745	13	46.4	1876	2	C70749	probable ppsa prot	818	12	42.9	54	2	B61492	ovomucoid (PSTI-cy
746	13	46.4	1959	1	AGRT	agrin - rat	819	12	42.9	54	2	A61492	ovomucoid (PSTI-cy
747	13	46.4	1964	2	T09059	notch4 - mouse	820	12	42.9	54	2	H31441	ovomucoid, thrid d
748	13	46.4	1978	2	T07081	acetyl-CoA carboxy	821	12	42.9	54	2	H31442	ovomucoid, thrid d
749	13	46.4	1984	2	T13171	probable vitelloge	822	12	42.9	54	2	H31442	ovomucoid, thrid d
750	13	46.4	1985	2	S19151	hypothetical prote	823	12	42.9	54	2	I31444	ovomucoid, thrid d
751	13	46.4	2004	2	T30185	hypothetical prote	824	12	42.9	54	2	B31444	ovomucoid, thrid d
752	13	46.4	2014	2	T21560	hypothetical prote	825	12	42.9	54	2	F31436	ovomucoid, thrid d
753	13	46.4	2039	2	S60123	hypothetical prote	826	12	42.9	54	2	G61494	ovomucoid (PSTI-cy
754	13	46.4	2039	2	S64540	probable calcium c	827	12	42.9	54	2	H31445	ovomucoid, thrid d
755	13	46.4	2056	2	G88564	protein R10B1.1 [l	828	12	42.9	54	2	F61494	ovomucoid (PSTI-cy
756	13	46.4	2108	2	H70819	probable polyketid	829	12	42.9	54	2	F31447	ovomucoid, thrid d
757	13	46.4	2124	2	T28619	polyketide synthas	830	12	42.9	54	2	F61492	ovomucoid (PSTI-cy
758	13	46.4	2126	2	E70522	probable polyketid	831	12	42.9	54	2	H61492	ovomucoid (PSTI-cy
759	13	46.4	2139	2	A44467	voltage-dependent	832	12	42.9	54	2	E31447	ovomucoid, thrid d

833	12	42.9	54	2	B31443	ovomucoid, third d	906	12	42.9	111	2	B85888	detox protein [imp
834	12	42.9	54	2	A31439	ovomucoid, third d	907	12	42.9	111	2	UC7528	guanylate cyclase
835	12	42.9	54	2	G61492	ovomucoid (PSI-ty	908	12	42.9	112	1	XIHU	collipase precursor
836	12	42.9	54	2	B31436	ovomucoid, third d	909	12	42.9	112	2	A46717	collipase precursor
837	12	42.9	54	2	B31436	ovomucoid, third d	910	12	42.9	112	2	S64448	hypothetical prote
838	12	42.9	55	2	B31442	ovomucoid, third d	911	12	42.9	113	2	P95248	conserved hypotet
839	12	42.9	56	2	D31445	ovomucoid, third d	912	12	42.9	113	2	C98113	conserved hypotet
840	12	42.9	56	2	A31446	ovomucoid, third d	913	12	42.9	115	2	S66824	probable membrane
841	12	42.9	56	2	C61588	ovomucoid (PSI-ty	914	12	42.9	115	2	G95059	hypothetical prote
842	12	42.9	56	2	G98233	hypothetical prote	915	12	42.9	115	2	B96664	probable RING zinc
843	12	42.9	56	2	A25356	bdellin B-3 - medi	916	12	42.9	115	2	S45920	hypothetical prote
844	12	42.9	59	2	A47218	carboxyl ester lip	917	12	42.9	119	1	PSN33W	phospholipase A2 (
845	12	42.9	60	2	T61886	zinc finger protei	918	12	42.9	121	1	QOEC13	hypothetical 13.5K
846	12	42.9	60	2	B86123	hypothetical prote	919	12	42.9	121	2	PH1661	Ig heavy chain V r
847	12	42.9	62	2	S62862	toxin III-8 - Tity	920	12	42.9	121	2	T02926	acyl carrier prote
848	12	42.9	62	2	S62866	toxin III-8 - Tity	921	12	42.9	121	2	A98089	hypothetical prote
849	12	42.9	62	2	B39510	toxin III-8 - Braz	922	12	42.9	121	2	C85934	hypothetical prote
850	12	42.9	62	2	S54336	metallothionein-2a	923	12	42.9	121	2	UQ1058	DNA-binding protei
851	12	42.9	62	2	S54335	metallothionein-2c	924	12	42.9	122	2	T18564	hypothetical prote
852	12	42.9	64	2	S66481	hypothetical prote	925	12	42.9	123	2	S30530	Ig heavy chain V r
853	12	42.9	66	2	E70083	hypothetical prote	926	12	42.9	124	2	T01251	hypothetical prote
854	12	42.9	67	2	JN0378	neurotoxin Os-3 -	927	12	42.9	125	1	NFB01	oxytocin / neuroph
855	12	42.9	67	2	S73050	probable acetyltra	928	12	42.9	125	1	NFRH1	oxytocin / neuroph
856	12	42.9	72	1	N2BPIJ	long neurotoxin 1	929	12	42.9	125	1	NFR01	oxytocin / neuroph
857	12	42.9	72	1	F82825	hypothetical prote	930	12	42.9	125	1	NFR1	oxytocin / neuroph
858	12	42.9	73	1	N2OH2	long neurotoxin 2	931	12	42.9	125	1	NFSH1	oxytocin / neuroph
859	12	42.9	77	2	B82617	hypothetical prote	932	12	42.9	125	2	A43755	oxytocin / neuroph
860	12	42.9	78	2	E70891	hypothetical prote	933	12	42.9	128	2	S26786	Ig heavy chain V r
861	12	42.9	82	2	S70807	hypothetical prote	934	12	42.9	128	2	UQ2180	hypothetical 14.8K
862	12	42.9	84	1	JU0152	acrosin/lypsin in	935	12	42.9	128	2	UQ1739	hypothetical 14.7K
863	12	42.9	85	2	D34770	ORF4 protein - sal	936	12	42.9	128	2	S27917	hypothetical prote
864	12	42.9	86	2	T61885	T3B protein - rabb	937	12	42.9	129	2	S01661	lysosomal (EC 3.2.1
865	12	42.9	89	2	D43692	proteinase inhibit	938	12	42.9	129	2	S07435	lysosomal (EC 3.2.1
866	12	42.9	89	2	S71555	wheat aluminum ind	939	12	42.9	130	2	G71212	hypothetical prote
867	12	42.9	89	2	UQ2361	protein mcl - Mont	940	12	42.9	131	2	T40656	hypothetical prote
868	12	42.9	92	2	T10783	probable chitinase	941	12	42.9	132	2	F64487	hypothetical prote
869	12	42.9	93	2	T06470	another-specific pr	942	12	42.9	133	2	T15955	hypothetical prote
870	12	42.9	94	2	T03285	collipase B precurs	943	12	42.9	133	2	T37311	hypothetical prote
871	12	42.9	96	1	XIHOB	keratin high-sulfu	944	12	42.9	133	2	JH0270	chondromodulin II
872	12	42.9	97	1	KRGTHM	keratin high-sulfu	945	12	42.9	133	2	JB0311	serine proteinase
873	12	42.9	97	1	KRSHM2	keratin high-sulfu	946	12	42.9	135	2	AF0865	conserved hypotet
874	12	42.9	97	2	T42335	hypothetical prote	947	12	42.9	135	2	S36197	hypothetical prote
875	12	42.9	97	2	JQ0041	hypothetical prote	948	12	42.9	135	2	G96691	hypothetical prote
876	12	42.9	97	2	D82789	hypothetical prote	949	12	42.9	136	2	T36241	lysosomal (EC 3.2.1
877	12	42.9	97	2	C95984	hypothetical prote	950	12	42.9	137	1	JC4233	lysosomal (EC 3.2.1
878	12	42.9	98	1	S25283	amoeapore A precu	951	12	42.9	137	2	T37213	hypothetical prote
879	12	42.9	98	1	KRSHM3	keratin high-sulfu	952	12	42.9	137	2	G84174	hypothetical prote
880	12	42.9	98	1	KRSHM4	keratin high-sulfu	953	12	42.9	138	2	AD0281	conserved hypotet
881	12	42.9	98	2	S26913	Ig heavy chain V r	954	12	42.9	138	2	S34725	regulation protein
882	12	42.9	98	2	I47083	BIIB2 high-sulfur	955	12	42.9	139	2	T23955	hypothetical prote
883	12	42.9	98	2	I47086	BIIB2 high-sulfur	956	12	42.9	140	2	JC5003	lysosome (EC 3.2.1
884	12	42.9	98	2	E96719	hypothetical prote	957	12	42.9	140	2	T49347	hypothetical prote
885	12	42.9	100	2	F95897	hypothetical prote	958	12	42.9	140	2	T05525	hypothetical prote
886	12	42.9	101	2	F75500	hypothetical prote	959	12	42.9	141	2	T33710	hypothetical prote
887	12	42.9	101	2	T26641	hypothetical prote	960	12	42.9	141	2	PC1294	trophozoite surfac
888	12	42.9	101	2	F72544	hypothetical prote	961	12	42.9	142	2	AC1934	protein B0294.1 [I
889	12	42.9	102	2	PH1249	Ig heavy chain V r	962	12	42.9	144	2	B89472	hypothetical prote
890	12	42.9	102	2	E89864	conserved hypotet	963	12	42.9	145	2	S74292	hemoglobin beta ch
891	12	42.9	102	2	S50530	hypothetical prote	964	12	42.9	145	2	S02027	F1511.18 [importe
892	12	42.9	102	2	A72604	neuropeptide Y pre	965	12	42.9	145	2	G96581	conserved hypotet
893	12	42.9	104	2	I50808	hypothetical prote	966	12	42.9	145	2	T34539	conserved hypotet
894	12	42.9	104	2	T43968	orf 61.2 - phage T	967	12	42.9	146	2	B82568	hypothetical prote
895	12	42.9	104	2	F45681	conserved hypotet	968	12	42.9	146	2	E95363	thyrotropin beta c
896	12	42.9	106	2	F72043	CT657 hypotetical	969	12	42.9	147	2	S34148	Ig heavy chain var
897	12	42.9	106	2	C86580	hypothetical prote	970	12	42.9	148	2	H71644	DNA-binding protei
898	12	42.9	107	2	T51038	conserved plasmid	971	12	42.9	148	2	B49727	hypothetical prote
899	12	42.9	107	2	A82642	hypothetical prote	972	12	42.9	148	2	T21334	helix-loop-helix p
900	12	42.9	108	2	E70931	conserved hypotet	973	12	42.9	149	2	JC5396	hypothetical prote
901	12	42.9	108	2	AD0715	aspergillopepsin I	974	12	42.9	151	2	T36314	keratin high-sulfu
902	12	42.9	109	2	JN0368	prepilin-like proc	975	12	42.9	152	1	KRSHMC	cysteine proteinas
903	12	42.9	109	2	S54446	detox protein [imp	976	12	42.9	152	2	S57427	high-sulfur wool m
904	12	42.9	111	2	G91043		977	12	42.9				
905	12	42.9	111	2	H65020		978	12	42.9				

152 2 147109 high-sulfur wool m  
12 42.9 152 2 147112 high-sulfur wool m  
980 12 42.9 152 2 C39384 finger protein Htr  
981 12 42.9 153 2 C70958 hypothetical prote  
982 12 42.9 153 2 A96751 hypothetical prote  
983 12 42.9 154 2 S36584 B6 protein - human  
984 12 42.9 154 2 A49727 DNA-binding protei  
985 12 42.9 154 2 S47524 gene id1 protein-  
986 12 42.9 154 2 T1ZM1 cryptin/factor XII  
987 12 42.9 155 1 JCS395 helix-loop-helix p  
988 12 42.9 155 2 JCS395 hypothetical prote  
989 12 42.9 155 2 JCS395 trypsin-plasmin in  
990 12 42.9 155 2 JCS395 trypsin-plasmin in  
991 12 42.9 157 2 S09805 vasococin / neurop  
992 12 42.9 158 2 B34332 hypothetical prote  
993 12 42.9 158 2 G86964 hypothetical prote  
994 12 42.9 158 2 T26444 hypothetical factor  
995 12 42.9 159 2 I84615 coagulation factor  
996 12 42.9 159 2 JQ0136 hypothetical 17.8k  
997 12 42.9 159 2 JQ0177 SalpR protein pre  
998 12 42.9 159 2 T33596 hypothetical prote  
999 12 42.9 160 1 ERAD45 early E3 18.5k gly  
1000 12 42.9 160 2 D48232 cysteine-rich exte

## ALIGNMENTS

RESULT 1  
GSFF3  
salivary glue protein sgs-3 - fruit fly (Drosophila melanogaster)  
C/Species: Drosophila melanogaster  
C/Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 09-Jul-2004  
C/Accession: A03329  
R/Garlinkel, M.D.; Prultt, R.E.; Meyerowitz, E.M.  
J. Mol. Biol. 169, 765-789, 1983  
A/Title: DNA sequences, gene regulation and modular protein evolution in the Drosophila  
A/Reference number: A92904; PMID:83294545; PMID:6411930  
A/Accession: A03329  
A/Molecule type: DNA  
A/Residues: 1-307 <GAR>  
A/Cross-references: UNIPROT:P02840; UNIPARC:UP100001245F4; GB:X01918; NID:98581; PIDN:CA  
C/Comment: This protein is produced by third-instar larvae.  
C/Genetics:  
A/Genes: sgs-3  
A/Cross-references: FlyBase:FBgn0003373  
A/Map position: 3L (68C)  
A/Intons: 10/1  
C/Superfamily: salivary glue protein  
C/Keywords: salivary gland; tandem repeat

Query Match 64.3%; Score 18; DB 1; Length 307;  
Best Local Similarity 16.7%; Pred. No. 1.3e-10;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 44 CTTTTTTTTTC 55

RESULT 2  
T24565  
hypothetical protein T06C12.14 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T24565  
R/Kelly, P.  
submitted to the EMBL Data Library, October 1996  
A/Reference number: Z19908  
A/Accession: T24565  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-245 <WLL>  
A/Cross-references: UNIPROT:O18039; UNIPARC:UP1000007C269; EMBL:Z81116; PIDN:CAB03300.1;

A/Experimental source: clone T06C12

C/Genetics:

A/Genes: CSP:T06C12.14

A/Map position: 5

A/Intons: 33/3; 121/1

C/Superfamily: Caenorhabditis elegans hypothetical protein C4967.3

Query Match 60.7%; Score 17; DB 2; Length 245;  
Best Local Similarity 16.7%; Pred. No. 5.8e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 193 CASATTTTSTC 204

RESULT 3  
D72630  
hypothetical protein APE1499 - Aeropyrum pernix (strain K1)

C/Species: Aeropyrum pernix  
C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C/Accession: D72630

R/Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A/Reference number: A72450; PMID:99310339; PMID:10382966

A/Accession: D72630

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-152 <KAW>

A/Cross-references: UNIPROT:Q9YBV0; UNIPARC:UP1000005DFPF; DBJ:AF000061; NID:95104821;

A/Experimental source: strain K1

C/Genetics:

A/Genes: APE1499

Query Match 57.1%; Score 16; DB 2; Length 152;  
Best Local Similarity 16.7%; Pred. No. 2.5e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 137 CSTATTTTSTRC 148

RESULT 4  
T31838  
hypothetical protein T05B4.12 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T31838

R/Bradshaw, H.

submitted to the EMBL Data Library, July 1997

A/Description: The sequence of C. elegans cosmid T05B4.

A/Reference number: Z21092

A/Accession: T31838

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-244 <BRA>

A/Cross-references: UNIPROT:O16421; UNIPARC:UP1000007ACTE; EMBL:AF016445; PIDN:AA69052

A/Experimental source: strain Bristol N2; clone T05B4

C/Genetics:

A/Genes: CSP:T05B4.12

A/Map position: 5

A/Intons: 30/3; 116/1

C/Superfamily: Caenorhabditis elegans hypothetical protein C4967.3

Query Match 57.1%; Score 16; DB 2; Length 244;  
Best Local Similarity 16.7%; Pred. No. 2.8e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12



Db 203 CTSYADSSSTSC 214

RESULT 5

TJ1841

hypothetical protein T05B4.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C:Accession: TJ1841

R:Bradshaw, H.

submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid T05B4.

A:Reference number: Z21092

A:Accession: TJ1841

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-248 <BR>

A:Cross-references: UNIPROT:O16424; UNIPARC:UPI00007F78C; EMBL:AF016445; PIDN:AAC69054.

A:Experimental source: strain Bristol N2; clone T05B4

C:Genetics:

A:Gene: CESP.T05B4.3

A:Map position: 5

A:Introns: 30/3; 118/1

C:Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3

Query Match 57.1%; Score 16; DB 2; Length 248;

Best Local Similarity 16.7%; Pred. No. 2.8e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 207 CTSFADSSSSC 218

RESULT 6

TJ1029

hypothetical protein F16H6.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: TJ1029

R:Matthews, U.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19362

A:Accession: TJ1029

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-257 <WIL>

A:Cross-references: UNIPROT:Q9XV90; UNIPARC:UPI000075F8F; EMBL:Z81506; PIDN:CAB04129.1;

A:Experimental source: clone F16H6

C:Genetics:

A:Gene: CESP.F16H6.3

A:Map position: 5

A:Introns: 39/3; 123/1

C:Superfamily: Caenorhabditis elegans hypothetical protein C49G7.3

Query Match 57.1%; Score 16; DB 2; Length 257;

Best Local Similarity 16.7%; Pred. No. 2.8e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 195 CTNATTTTQASC 206

RESULT 7

T52565

glutathione synthase (EC 6.3.2.3) [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 09-Jul-2004

C:Accession: T52565

R:Skipsey, M.; Andrews, C.J.; Townsend, J.K.; Jepson, I.; Edwards, R.

submitted to the EMBL Data Library, July 1999

A:Description: Isolation of cDNA and genomic clones of glutathione synthetase containing

A:Reference number: Z26117

A:Accession: T52565

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-539 <SKI>

A:Cross-references: UNIPROT:P46416; UNIPARC:UPI000016DB47; EMBL:AJ243813; PIDN:CAB51027

A:Experimental source: cultivar Columbia; seedling; whole plant

C:Genetics:

A:Gene: gsh2

C:Superfamily: glutathione synthase

C:Keywords: ligase

Query Match 57.1%; Score 16; DB 2; Length 539;

Best Local Similarity 16.7%; Pred. No. 3.2e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 5 CSLSYSSSSTC 16

RESULT 8

A48579

trichozoite surface protein TSP11 - Giardia lamblia

C:Species: Giardia lamblia

C>Date: 01-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A48579

R:By, P.L.; Khanna, K.K.; Manning, P.A.; Mayrhofer, G.

MoI. Biochem. Parasitol. 58, 247-257, 1993

A:Title: A gene encoding a 69-kilodalton major surface protein of Giardia intestinalis

A:Reference number: A48579; PMID:8479449

A:Contents: Ad-1

A:Accession: A48579

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-667 <EY1>

A:Cross-references: UNIPROT:Q03185; UNIPARC:UPI0000137726; GB:M95814; NID:g159106; PID:

A:Note: sequence extracted from NCBI backbone (NCBIN:130056, NCBIIP:130058)

Query Match 57.1%; Score 16; DB 2; Length 667;

Best Local Similarity 16.7%; Pred. No. 3.4e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 449 CSTRARTVATC 460

RESULT 9

T51024

related to C2H2 zinc finger transcription factor D-Spl [imported] - Neurospora crassa

N:Alternate names: protein B7F21.50

C:Species: Neurospora crassa

C>Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 09-Jul-2004

C:Accession: T51024

R:Schulte, U.; Aign, V.; Hohensei, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, July 2000

A:Reference number: Z25286

A:Accession: T51024

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-770 <SCH>

A:Cross-references: UNIPROT:Q9P119; UNIPARC:UPI000017B4E0; EMBL:AL389901; GSPDB:GN00116,

A:Experimental source: BAC clone B7F21; strain OR74A

C:Genetics:

A:Gene: NCSB:B7F21.50

A:Map position: 6

A:Introns: 117/1

Query Match 57.1%; Score 16; DB 2; Length 770;

Best Local Similarity 16.7%; Pred. No. 3.5e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
Db 92 CSATTAAVGTAC 103

## RESULT 10

T31097  
chitin synthase (EC 2.4.1.16) CHS1 - fungus (*Filobasidium floriforme*)  
C/Species: *Filobasidium neoformans*; *Cryptococcus neoformans*  
C/Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T31097  
R/Species: C.A.  
submitted to the EMBL Data Library, August 1997

A/Description: CHS1, a class IV chitin synthase of *Cryptococcus neoformans*.  
A/Reference number: Z20980  
A/Accession: T31097  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-1041 <SPS>  
A/Cross-references: UNIPROT:O13356; UNIPARC:UPI00001278CD; EMBL:AF021318; NID:g2444456;  
A/Experimental source: strain H99  
C/Genetics:  
A/Gene: CHS1  
A/Introns: 152/1; 556/1; 634/3; 922/2  
C/Function:  
A/Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucose  
A/Pathway: chitin biosynthesis  
A/Note: class IV chitin synthase  
C/Superfamily: chitin synthase chs4  
C/Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 57.1%; Score 16; DB 2; Length 1041;  
Best Local Similarity 16.7%; Pred. No. 3.7e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
Db 212 CNSTSSSSGTC 223

## RESULT 11

AC0416  
probable prelinin peptidase dependent protein D precursor ppdd [imported] - *Yersinia pestis*  
C/Species: *Yersinia pestis*  
C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C/Accession: AC0416  
R/Parthill, J.; Wren, B.W.; Thomson, N.R.; Tilball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.; 11. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001  
A/Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A/Reference number: AB0001; MUID:21470413; PMID:11586360  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-154 <RUR>  
A/Cross-references: UNIPROT:Q8ZB15; UNIPARC:UPI000000D98A; GB:AL590842; PIDN:CAC92655.1;  
C/Genetics:  
A/Gene: ppdd  
C/Superfamily: gonococcal fimbrial protein

Query Match 53.6%; Score 15; DB 2; Length 154;  
Best Local Similarity 16.7%; Pred. No. 1.2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
Db 129 CTATDSAMDSC 140

## RESULT 12

B41378  
cytochrome c553i precursor - *Paracoccus denitrificans*

C/Species: *Paracoccus denitrificans*  
C/Date: 28-May-1992 #sequence\_revision 28-May-1992 #text\_change 09-Jul-2004  
C/Accession: B41378  
R/Ras, J.; Reijnders, W.N.M.; Van Spanning, R.J.M.; Harms, N.; Oltmann, L.F.; Stouthamer J. Bacteriol. 173, 6971-6979, 1991

A/Title: Isolation, sequencing, and mutagenesis of the gene encoding cytochrome c-553i o  
A/Reference number: A41378; MUID:92041583; PMID:1657873  
A/Accession: B41378  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-226 <RAS>  
A/Cross-references: UNIPROT:P29967; UNIPARC:UPI0000126C5E; GB:M75583; NID:g150576; PIDN:  
C/Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein  
F/1-20/Domain: signal sequence #status predicted <SIG>  
F/1-226/Product: cytochrome c553i #status predicted <MAT>  
F/115-193/Domain: cytochrome c6 homology <CYC>  
F/125,128/Binding site: heme (Cys) (covalent) #status predicted  
F/129,173/Binding site: heme iron (His, Met) (axial ligands) #status predicted

Query Match 53.6%; Score 15; DB 2; Length 226;  
Best Local Similarity 16.7%; Pred. No. 1.3e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
Db 16 CAASAIAGTALC 27

## RESULT 13

T21786  
hypothetical protein F35E8.9 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T21786  
R/Gardner, A.  
submitted to the EMBL Data Library, November 1996  
A/Reference number: Z19472  
A/Accession: T21786  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-248 <WIL>  
A/Cross-references: UNIPROT:Q45450; UNIPARC:UPI0000077E4A; EMBL:281529; PIDN:CAN04292.1,  
A/Experimental source: clone F35E8  
C/Genetics:  
A/Gene: CSR.F35E8.9  
A/Map position: 5  
A/Introns: 33/3; 121/1  
C/Superfamily: *Caenorhabditis elegans* hypothetical protein C49G7.3

Query Match 53.6%; Score 15; DB 2; Length 248;  
Best Local Similarity 16.7%; Pred. No. 1.3e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
Db 207 CSSYIADSTAC 218

## RESULT 14

A84746  
probable NAM (no apical meristem)-like protein [imported] - *Arabidopsis thaliana*  
C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: A84746  
R/Alin, X.; Kaul, S.; Rounsley, S.D.; Snea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L. eues, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salberg, S.L.; Frazer, C.M.; Venter, J Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: A84746  
A/Status: preliminary  
A/Molecule type: DNA

A:Residues: 1-268 <STO>  
 A:Cross-references: UNIPROT:Q22798; UNIPARC:UPI00000A7D54; GB:AE002093; NID:92459430; P1  
 C:Genetics:  
 A:Gene: At2g33480  
 A:Map position: 2

Query Match 53.6%; Score 15; DB 2; Length 268;  
 Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
 DB 230 CCSSTSTSSSSVC 241

RESULT 15  
 S69193  
 Probable finger protein Pazf1 - garden pea  
 C:Species: Pisum sativum (garden pea)  
 C>Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C:Accession: S69193  
 R:Michael, A.J.; Hofer, J.M.I.; Ellis, T.H.N.  
 Plant Mol. Biol. 30, 1051-1058, 1996  
 A>Title: Isolation by PCR of a cDNA clone from pea petals with similarity to petunia and  
 A:Reference number: S69193; MUID:96270382; PMID:8639742  
 A:Accession: S69193  
 A:Molecule type: mRNA  
 A:Residues: 1-273 <MIC>  
 A:Cross-references: UNIPROT:Q41070; UNIPARC:UPI000009E1A7; EMBL:X87374; NID:9854399; P1D  
 C:Superfamily: Arabidopsis thaliana hypothetical protein F12E4.290  
 C:Keywords: DNA binding; zinc finger  
 F:115-135/Region: zinc finger CCH motif  
 F:202-222/Region: zinc finger CCH motif

Query Match 53.6%; Score 15; DB 2; Length 273;  
 Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
 DB 230 CTVANTTVTSSC 40

RESULT 16  
 AH2744  
 Conserved hypothetical protein Atu1368 [imported] - Agrobacterium tumefaciens (strain C5  
 C:Species: Agrobacterium tumefaciens  
 C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
 C:Accession: AH2744  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Author: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB5577; MUID:21608550; PMID:11743193  
 A:Accession: AH2744  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-283 <KUR>  
 A:Cross-references: UNIPROT:Q8UFW9; UNIPARC:UPI00001645DD; GB:AE008688; PIDN:AAL42374.1;  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu1368  
 A:Map position: circular chromosome

Query Match 53.6%; Score 15; DB 2; Length 283;  
 Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 242 CSATHSMDMAC 253

RESULT 17  
 S57770  
 xyloglucan endo-1,4-beta-D-glucanase (EC 3.2.1.-) precursor (clone XET-B2) - tomato  
 C:Species: Lycopersicon esculentum (tomato)  
 C>Date: 06-Dec-1996 #sequence\_revision 07-Feb-1997 #text\_change 09-Jul-2004  
 C:Accession: S57770; S49813  
 R:Arrowsmith, D.A.; de Silva, J.  
 Plant Mol. Biol. 28, 391-403, 1995  
 A>Title: Characterisation of two tomato fruit-expressed cDNAs encoding xyloglucan endo-1  
 A:Reference number: S57769; MUID:95359399; PMID:7632911  
 A:Accession: S57770  
 A>Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-287 <ARR>  
 A:Cross-references: UNIPROT:Q43528; UNIPARC:UPI00000AAE06; EMBL:X82684; NID:9577067; P1  
 A>Note: the sequence from Fig. 2 is inconsistent with that from Fig. 5 in having 246-Gl  
 C:Superfamily: endoxyloglucan transferase  
 C:Keywords: glycosidase; hydrolase  
 F:1-18/Domain: signal sequence  
 F:19-287/Product: xyloglucan endo-transglycosylase #status predicted <STIG>

Query Match 53.6%; Score 15; DB 2; Length 287;  
 Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
 DB 217 CWSAASATSTSSC 228

RESULT 18  
 S49812  
 xyloglucan endo-1,4-beta-D-glucanase (EC 3.2.1.-) precursor (clone XET-B1) - tomato  
 C:Species: Lycopersicon esculentum (tomato)  
 C>Date: 05-Mar-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
 C:Accession: S57769; S49812  
 R:Arrowsmith, D.A.; de Silva, J.  
 Plant Mol. Biol. 28, 391-403, 1995  
 A>Title: Characterisation of two tomato fruit-expressed cDNAs encoding xyloglucan endo-1  
 A:Reference number: S57769; MUID:95359399; PMID:7632911  
 A:Accession: S57769  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-289 <AR2>  
 A:Cross-references: UNIPROT:Q43527; UNIPARC:UPI00000A4B68; EMBL:X82685; NID:9577065; P1T  
 C:Genetics:  
 A:Gene: XET-B1  
 C:Superfamily: endoxyloglucan transferase  
 C:Keywords: glycosidase; hydrolase

Query Match 53.6%; Score 15; DB 2; Length 289;  
 Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
 DB 219 CWSAASATSTSSC 230

RESULT 19  
 G97525  
 CG10208 protein (AE003744) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)  
 C:Species: Agrobacterium tumefaciens  
 C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
 C:Accession: G97525  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: G97525  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-318 <KUR>  
C:Cross-references: UNIPROT:Q8UFW9; UNIPARC:UPI00000D1B27; GB:AE007869; P1DN:AAK87160.1;  
C:Genetics:  
A:Gene: AGR\_C\_2528  
A:Map position: circular chromosome

Query Match 53.6%; Score 15; DB 2; Length 318;  
Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 277 CSATHTSDMAAC 288

RESULT 20  
T21152  
hypothetical protein F20E11.7 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T21152  
R:Ainscough, R.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19383  
A:Accession: T21152  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-357 <MIT>  
A:Cross-references: UNIPROT:Q9XV67; UNIPARC:UPI0000060FAD; EMBL:Z81508; P1DN:CAE04138.1;  
A:Experimental source: clone F20E11  
C:Genetics:  
A:Gene: CESP:F20E11.7  
A:Map position: 5  
A:introns: 52/3; 88/2; 103/3; 154/3; 190/2; 205/3; 241/2; 270/3; 321/3

Query Match 53.6%; Score 15; DB 2; Length 357;  
Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 7 CTATTGTMSAGC 18

RESULT 21  
S18765  
Sd protein - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C>Date: 19-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S18765  
R:Powers, P.A.; Ganetzky, B.  
Genetics 129, 133-144, 1991  
A:Title: On the components of segregation distortion in *Drosophila melanogaster*. V. Molecular  
A:Reference number: S18765; MUID:92038937; PMID:1936954  
A:Molecule type: mRNA  
A:Residues: 1-363 <POW>  
C:Cross-references: UNIPROT:P25722; UNIPARC:UPI000016BD5D; EMBL:X60218; NID:G11012; P1DN:  
C:Genetics:  
A:Gene: FlyBase:Sd  
A:Cross-references: FlyBase:FBgn0024230

Query Match 53.6%; Score 15; DB 2; Length 363;  
Best Local Similarity 16.7%; Pred. No. 1.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 331 CSRTSTPNSC 342

RESULT 22  
T27303  
hypothetical protein Y68A4A.10 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T27303  
R:Steward, C.  
submitted to the EMBL Data Library, January 1998  
A:Reference number: Z20340  
A:Accession: T27303  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-392 <MIT>  
A:Cross-references: UNIPROT:Q9XXP6; UNIPARC:UPI000007BED1; EMBL:AL021503; P1DN:CAA16424.  
A:Experimental source: clone Y68A4A  
C:Genetics:  
A:Gene: CESP:Y68A4A.10  
A:Map position: 5  
A:introns: 45/2; 74/3; 139/3; 175/2; 190/3; 241/3; 305/3; 356/3

Query Match 53.6%; Score 15; DB 2; Length 392;  
Best Local Similarity 16.7%; Pred. No. 1.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 260 CTATTGTTSTGC 271

RESULT 23  
T08928  
sucrose cleavage protein homolog T15N24.70 - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: T08928  
R:Bevan, M.; Zimmermann, W.; Gruenisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May  
submitted to the Protein Sequence Database, May 1999  
A:Reference number: Z16518  
A:Accession: T08928  
A:Molecule type: DNA  
A:Residues: 1-409 <BEV>  
A:Cross-references: UNIPROT:Q9SUA2; UNIPARC:UPI000009E309; EMBL:AL078465; GSPDB:GN00062,  
A:Experimental source: cultivar Columbia; BAC clone T15N24  
C:Genetics:  
A:Gene: ATSP:T15N24.70  
A:Map position: 4  
A:introns: 126/3; 155/2; 223/1; 299/3

Query Match 53.6%; Score 15; DB 2; Length 409;  
Best Local Similarity 16.7%; Pred. No. 1.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 394 CAAATVAVAYTC 405

RESULT 24  
S74388  
gamma-glutamyl phosphate reductase proA - *Synechocystis* sp. (strain PCC 6803)  
N:Alternate names: hypothetical protein sl10461  
C:Species: *Synechocystis* sp.  
A:Variety: PCC 6803  
C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: S74388  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
s.  
A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S74388  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-420 <RAN>  
A:Cross-references: UNIPROT:Q55167; UNIPARC:UPI000013848; EMBL:U64001; GB:AB001339; NID  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: PROA  
C:Superfamily: gamma-glutamyl phosphate reductase

Query Match 53.6%; Score 15; DB 2; Length 420;  
Best Local Similarity 16.7%; Pred. No. 1.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 153 CGASHSTAAIC 164

RESULT 25  
D96552  
unknown protein, 74043-75895 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: D96552  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Hlitzar, L.  
Nature 408, 616-620, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luvo, J.S.; Maltl, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: D96552  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-423 <STO>  
A:Cross-references: UNIPROT:Q9C8U4; UNIPARC:UPI00000A997F; GB:AE05173; NID:g10092345; F  
C:Genetics:  
A:Gene: FSD21.9  
A:Map position: 1  
C:Superfamily: Arabidopsis thaliana hypothetical protein F2AB22.230

Query Match 53.6%; Score 15; DB 2; Length 423;  
Best Local Similarity 16.7%; Pred. No. 1.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 306 CTTTSSLKASC 317

RESULT 26  
S31126  
hypothetical protein F59B2.6 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S31126  
R:Sulston, J.; Du, Z.; Thomas, K.; Wilson, R.; Hillier, L.; Staden, R.; Halloran, N.; Gh  
awkins, T.; Ainscough, R.; Waterston, R.  
submitted to the EMBL Data Library, November 1991  
A:Description: The C. elegans sequencing project: A beginning.  
A:Reference number: S31122  
A:Accession: S31126  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-503 <SUL>  
A:Cross-references: UNIPROT:P34482; UNIPARC:UPI000013B9D; EMBL:Z11505; NID:g6718; PID:g  
C:Genetics:  
A:introns: 26/3; 42/2; 121/3; 367/3; 443/1

C:Superfamily: Caenorhabditis elegans hypothetical protein F59B2.6

Query Match 53.6%; Score 15; DB 2; Length 503;  
Best Local Similarity 16.7%; Pred. No. 1.5e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 4 CSASTSQLSTPFC 15

RESULT 27  
B87180  
pyruvate (or indolepyruvate) decarboxylase [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 05-Oct-2004  
C:Accession: B87180  
R:Coile, S.T.; Eigimeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H  
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Fellwell, T.; Fraser, A.; Hamlin, N.; Holroyd  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S  
A:Title: Massive gene decay in the leprosy bacillus.  
A:Reference number: A86909; MUID:21128732; PMID:11234002  
A:Accession: B87180  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-569 <STO>  
A:Cross-references: UNIPROT:Q9CDB6; UNIPARC:UPI00000C6E3D; GB:AL450380; NID:g13093715;  
C:Genetics:  
A:Gene: pdc  
C:Superfamily: pyruvate decarboxylase/indolepyruvate decarboxylase; thiamin pyrophospha

Query Match 53.6%; Score 15; DB 2; Length 569;  
Best Local Similarity 16.7%; Pred. No. 1.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 140 CSQANLTTATAC 151

RESULT 28  
JC7818  
metalloprotease I precursor - Alteromonas sp. (Strain O-7)  
C:Species: Alteromonas sp. (Strain O-7)  
C>Date: 03-Jun-2002 #sequence\_revision 03-Jun-2002 #text\_change 31-Dec-2004  
C:Accession: JC7818  
R:Myamoto, K.; Tsujibo, H.; Nukui, E.; Itoh, H.; Kaiznu, Y.; Inamori, Y.  
Bioosci. Biotechnol. Biochem. 66, 416-421, 2002  
A:Title: Isolation and characterization of the genes encoding two metalloproteases (Mpr  
A:Reference number: JC7818; PMID:11999419; MUID:21294062  
A:Accession: JC7818  
A:Molecule type: DNA  
A:Residues: 1-727 <MTY>  
A:Cross-references: UNIPROT:Q8VLQ5; UNIPARC:UPI000017CD49; DDBJ:AB063611  
A:Experimental source: strain O-7  
C:Comment: This enzyme, an extracellular alkaline metalloprotease, showing high similar  
C:Genetics:  
A:Gene: mpr1  
A:Start codon: GTG  
C:Superfamily: zinc metalloendopeptidase, neutral protease type (elastase)

Query Match 53.6%; Score 15; DB 2; Length 727;  
Best Local Similarity 16.7%; Pred. No. 1.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 581 CNSTSTSNESC 592

RESULT 29

S40382  
box A-binding factor - fruit fly (*Drosophila melanogaster*)  
N:Alternate names: ABF; transcription factor GATAB  
C:Species: *Drosophila melanogaster*  
C>Date: 13-Jan-1995 #sequence\_revision 06-Sep-1996 #text\_change 09-Jul-2004  
C/Accession: S40382  
R:Abel, T.; Michelson, A.M.; Maniatis, T.  
Development 119, 623-633, 1993  
A>Title: A *Drosophila* GATA family member that binds to Adh regulatory sequences is expressed in the developing embryo  
A:Reference number: S40382; UID:9424445; PMID:8187633  
A:Accession: S40382  
A:Molecule type: mRNA  
A:Residues: 1-779 <ABF>  
A:Cross-references: UNIPROT:P52172; UNIPARC:UPI0000135F5F; EMBL:X76217; NID:9441491; PID:9441491  
C:Comment: This transcriptional activator is the earliest known marker of the developing embryo  
C:Genetics:  
A:Gene: FlyBase:FBgn0003507  
A:Cross-references: FlyBase:FBgn0003507  
C:Superfamily: box A-binding factor; GATA-type zinc finger homology  
C:Keywords: DNA binding; nucleus; transcription factor; transcription regulation; zinc finger  
F:316-369/Domain: GATA-type zinc finger homology <GZF>  
F:319-343/Region: zinc finger GATA motif

Query Match 53.6%; Score 15; DB 1; Length 779;  
Best Local Similarity 16.7%; Pred. No. 1.7e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 507 C5S1SSNRAC 518

## RESULT 30

PRVZAM  
spheroidin precursor - *Amsacta moorei* poxvirus  
C:Species: *Amsacta moorei* poxvirus  
C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
C/Accession: JQ1436; P00289; E41561  
R:Banville, M.; Dumas, F.; Trifiro, S.; Arif, B.; Richardson, C.  
J. Gen. Virol. 73, 559-566, 1992  
A>Title: The predicted amino acid sequence of the spheroidin protein from *Amsacta moorei*  
A:Reference number: JQ1436; UID:92185464; PMID:1545219  
A:Accession: JQ1436  
A:Molecule type: DNA  
A:Residues: 1-1003 <BAN>  
A:Cross-references: UNIPROT:P29815; UNIPARC:UPI00000008A1; GB:M75889; NID:9209629; PID:9209629  
A:Accession: P00289  
A:Molecule type: protein  
A:Residues: 2-7, 'X', 9-11, 'X', 13-86-90, 'X', 92-98, 'X', 100-102;532-540;728-750;786-802;810-812  
A:Cross-references: UNIPARC:UPI00001749B4; UNIPARC:UPI00001749B5; UNIPARC:UPI00001749B6;  
R:Hall, R.L.; Moyer, R.W.  
J. Virol. 65, 6516-6527, 1991  
A>Title: Identification, cloning, and sequencing of a fragment of *Amsacta moorei* entomopoxvirus  
A:Reference number: A41561; UID:92046310; PMID:1942245  
A:Accession: E41561  
A:Molecule type: DNA  
A:Residues: 1-1003 <HAI>  
A:Cross-references: UNIPARC:UPI00000008A1; GB:M77182; NID:9209631; PID:AAA42383.1; PID:AAA42383.1  
C:Comment: This protein is a major component of the occlusion body which serves to protect the virus  
C:Superfamily: *Amsacta moorei* poxvirus spheroidin  
C:Keywords: acetylated amino end; glycoprotein; leucine zipper  
F:2-1003/Product: spheroidin #status experimental <MAT>  
F:119-140/Region: leucine zipper motif  
F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental  
F:47,109,204,210,259,510,559,547,561,799,901,922,955/Binding site: carbohydrate (Asn) (C)  
F:537,799/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 53.6%; Score 15; DB 1; Length 1003;  
Best Local Similarity 16.7%; Pred. No. 1.8e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 951 CADANSSSDSC 962

RESULT 31  
T19734  
hypothetical protein C34F6.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T19734  
R:White, S.  
Submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19171  
A:Accession: T19734  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1043 <WIL>  
A:Cross-references: UNIPROT:O17644; UNIPARC:UPI00000775AC; EMBL:Z81479; PID:CA03944.1;  
A:Experimental source: clone C34F6  
C:Genetics:  
A:Gene: CBSP:C34F6.1  
A:Map position: X  
A:Introns: 17/3; 62/1; 129/3; 181/1; 292/1; 348/1; 404/1; 665/1; 769/1; 876/1; 977/3

Query Match 53.6%; Score 15; DB 2; Length 1043;  
Best Local Similarity 16.7%; Pred. No. 1.8e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 218 CHTGASFRTAC 229

## RESULT 32

T18210  
delta endotoxin - *Bacillus thuringiensis*  
C:Species: *Bacillus thuringiensis*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 05-Oct-2004  
C/Accession: T18210  
R:Narva, K.E.; Payne, J.; Feitelson, J.S.  
Submitted to the EMBL Data Library, August 1994  
A:Description: Novel *Bacillus thuringiensis* toxins active against coleopteran larvae.  
A:Reference number: Z18826  
A:Accession: T18210  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1186 <NAR>  
A:Cross-references: UNIPROT:Q45710; UNIPARC:UPI0000127491; EMBL:U13955; NID:9538377; PID:9538377  
C:Keywords: delta-endotoxin

Query Match 53.6%; Score 15; DB 2; Length 1186;  
Best Local Similarity 16.7%; Pred. No. 1.8e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 466 CTTSTANSYKASC 477

## RESULT 33

T10236  
xanthine dehydrogenase homolog T1111.140 - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C/Accession: T10236  
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Jesse, T.; Heijnen, L.; Vos, P.; Mewes, F.  
Submitted to the Protein Sequence Database, June 1999  
A:Reference number: Z16992  
A:Accession: T10236  
A:Molecule type: DNA  
A:Residues: 1-1364 <BEV>  
A:Cross-references: UNIPROT:Q9SW45; UNIPARC:UPI000009E008; EMBL:AL079347; GSPDB:GN00062;  
A:Experimental source: cultivar Columbia; BAC clone T1111

Query Match 53.6%; Score 15; DB 1; Length 1364;  
Best Local Similarity 16.7%; Pred. No. 1.8e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

A:Gene: ATSP:T1111.140  
A:Map position: 4  
A:Introns: 35/1; 66/2; 115/3; 613/3; 694/3; 799/3; 879/3; 974/3; 1048/3; 1102/3; 1147/3;  
C:Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology  
C:Keywords: 2Fe-2S; metalloprotein  
F:43; 48; 51; 86/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 53.6%; Score 15; DB 2; Length 1364;  
Best Local Similarity 16.7%; Pred. No. 1.9e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 209 CGSKTTSEATC 220

RESULT 34  
S65571  
pattern formation protein GNOM - Arabidopsis thaliana  
N:Alternate names: EMB30 protein  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 09-Jul-2004  
C:Accession: S65571; S65572  
Mol. Gen. Genet. 250, 681-691, 1996  
R:Busch, M.; Mayer, U.; Uerogens, G.  
A:Title: Molecular analysis of the Arabidopsis pattern formation gene GNOM: gene structure  
A:Reference number: S65571; MUID:96204508; PMID:8628228  
A:Accession: S65571  
A:Molecule type: DNA  
A:Residues: 1-1451 <BUS>  
A:Cross-references: UNIPROT:Q42510; UNIPARC:UPI000012982; EMBL:U56433; NID:G1209632; PI  
A:Accession: S65572  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-110; 112-866; 'G', 868-1451 <BUS>  
A:Cross-references: UNIPARC:UPI000016DA5; EMBL:U56432; NID:G1209630; PID:G1209631  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995  
C:Genetics:  
A:Gene: GNOM; EMB30  
A:Introns: 246/3

Query Match 53.6%; Score 15; DB 2; Length 1451;  
Best Local Similarity 16.7%; Pred. No. 1.9e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 22 CTDSNTTTLAC 33

RESULT 35  
T25888  
hypothetical protein T10H10.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T25888  
R:Neison, J.; Langston, Y.  
Submitted to the EMBL Data Library, December 1996  
A:Description: The sequence of C. elegans cosmid T10H10.  
A:Reference number: Z20105  
A:Accession: T25888  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2098 <NEU>  
A:Cross-references: UNIPROT:P91443; UNIPARC:UPI00000756CE; EMBL:U80848; PIDN:AA837988.1;  
A:Experimental source: strain Bristol N2; clone T10H10  
C:Genetics:  
A:Gene: CESP:T10H10.1  
A:Map position: X  
A:Introns: 6/3; 43/3; 92/3; 154/2; 226/1; 280/3; 417/2; 472/3; 812/3; 853/3; 887/3; 978/  
/3  
F:65-720/Domain: myosin motor domain homology <MMO>

Query Match 53.6%; Score 15; DB 2; Length 2098;  
Best Local Similarity 16.7%; Pred. No. 2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 1188 CADSATTAELC 1199

RESULT 36  
T13806  
toucan gene protein - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: T13806  
R:Grammont, M.; Daatague, B.; Couderc, J.L.  
Development 124, 4917-4926, 1997  
A:Title: The Drosophila toucan (toc) gene is required in germline cells for somatic cell  
A:Reference number: Z17769; MUID:98090047; PMID:9362455  
A:Accession: T13806  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-2176 <GBA>  
A:Cross-references: UNIPROT:Q46112; UNIPARC:UPI0000083F2C; EMBL:Y14157; NID:G2760521; P  
A:Genetics:  
A:Gene: toc  
A:Cross-references: FlyBase:FBgn0015600

Query Match 53.6%; Score 15; DB 2; Length 2176;  
Best Local Similarity 16.7%; Pred. No. 2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 2031 CTKASATTSATC 2042

RESULT 37  
S60224  
polyketide synthase 1 - Colletotrichum lagenarium  
C:Species: Colletotrichum lagenarium  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S60224  
R:Takano, Y.; Kubo, Y.; Shimizu, K.; Mise, K.; Okuno, T.; Furusawa, I.  
Mol. Gen. Genet. 249, 162-167, 1995  
A:Title: Structural analysis of PKS1, a polyketide synthase gene involved in melanin biosynthesis  
A:Reference number: S60224; MUID:96086932; PMID:7500937  
A:Accession: S60224  
A>Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-2187 <TKA>  
A:Cross-references: UNIPROT:P79068; UNIPARC:UPI000006C308; GB:D83643; NID:G1208941; PIDN  
A:Genetics:  
A:Gene: PKS1  
A:Function:  
A:Pathway: melanin biosynthesis  
C:Keywords: carrier protein  
F:402-811/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>  
F:910-1198/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>  
F:166-1736/Domain: acyl carrier protein homology <ACP>  
F:1788-1862/Domain: acyl carrier protein homology <ACPI>

Query Match 53.6%; Score 15; DB 2; Length 2187;  
Best Local Similarity 16.7%; Pred. No. 2e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 553 CSSAAALNVAC 564

RESULT 38  
T28675

alpha-51D immobilization antigen - Paramedium tetraurelia  
C:Species: Paramedium tetraurelia  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T28675  
R:Schwegmann, K.J.  
submitted to the EMBL Data Library, March 1996  
A:Reference number: Z20506  
A:Accession: T28675  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2533 <SCH>  
C:Cross-references: UNIPROT:P90589; UNIPARC:UPI0000078A73; EMBL:X96400; PIDN:CAA65264.1  
A:Gene: alpha-51D  
A:Genetic code: SGCS  
A:introns: 280/3; 538/2; 1248/2  
C:Superfamily: G surface protein

Query Match 53.6%; Score 15; DB 2; Length 2533;  
Best Local Similarity 16.7%; Pred. No. 2.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 1 CXXXXXXXXXXC 12  
Db 360 CKGTSTTEAAC 371

RESULT 39  
T28674  
alpha-51D-immobilization antigen - Paramedium tetraurelia  
C:Species: Paramedium tetraurelia  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T28674  
R:Schmidt, H.J.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: Z20505  
A:Accession: T28674  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2533 <SCH>  
C:Cross-references: UNIPROT:Q27183; UNIPARC:UPI000007E3B2; EMBL:X85135; NID:G728634; PIDN:G728634  
C:Genetics:  
A:Genetic code: SGCS  
A>Note: alpha-51D  
C:Superfamily: G surface protein

Query Match 53.6%; Score 15; DB 2; Length 2533;  
Best Local Similarity 16.7%; Pred. No. 2.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 1 CXXXXXXXXXXC 12  
Db 360 CKGTSTTEAAC 371

RESULT 40  
T31687  
surface antigen - Paramedium primaurelia  
C:Species: Paramedium primaurelia  
C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C:Accession: T31687  
R:Bourgain-Guglielmi, F.; Caron, F.  
Journal of Bacteriol. Microbiol. 43, 303-314, 1996  
A:Title: Molecular characterization of the D surface protein gene subfamily in Paramedium primaurelia  
A:Reference number: Z21061; MUID:96313351; PMID:8768434  
A:Accession: T31687  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2543 <BOU>  
C:Cross-references: UNIPROT:P90649; UNIPARC:UPI000007A903; EMBL:X96616; NID:G1235576; PIDN:G1235576  
A:Genetic code: SGCS  
C:Superfamily: G surface protein

Query Match 53.6%; Score 15; DB 2; Length 2543;  
Best Local Similarity 16.7%; Pred. No. 2.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 1 CXXXXXXXXXXC 12  
Db 360 CKGTSTTEAAC 371

RESULT 41  
A24420  
notch protein - fruit fly (Drosophila melanogaster)  
N:Alternate names: neurogenic repetitive locus protein  
C:Species: Drosophila melanogaster  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 05-Oct-2004  
C:Accession: A24420; A24768; S09358; A05267  
R:Kidd, S.; Kelley, M.R.; Young, M.W.  
Mol. Cell. Biol. 6, 3094-3108, 1986  
A:Reference number: A24420; MUID:87064624; PMID:3097517  
A:Accession: A24420  
A:Molecule type: DNA  
A:Residues: 1-2703 <KID>  
C:Cross-references: UNIPROT:P07207; UNIPARC:UPI000016BCC6; GB:K03508; NID:G157991; PIDN:G157991  
R:Wharton, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S.  
Cell 43, 567-581, 1985  
A:Reference number: A24768; MUID:86079539; PMID:3935325  
A:Accession: A24768  
A:Molecule type: mRNA  
A:Residues: 1-48, 'I', '50-118, 'R', '120-230, 'I', '232-256, 'N', '258-266, 'A', '268-872, 'R', '874-958, 'R', '959-1000, 'I', '1001-1002, 'I', '1003-1004, 'I', '1005-1006, 'I', '1007-1008, 'I', '1009-1010, 'I', '1011-1012, 'I', '1013-1014, 'I', '1015-1016, 'I', '1017-1018, 'I', '1019-1020, 'I', '1021-1022, 'I', '1023-1024, 'I', '1025-1026, 'I', '1027-1028, 'I', '1029-1030, 'I', '1031-1032, 'I', '1033-1034, 'I', '1035-1036, 'I', '1037-1038, 'I', '1039-1040, 'I', '1041-1042, 'I', '1043-1044, 'I', '1045-1046, 'I', '1047-1048, 'I', '1049-1050, 'I', '1051-1052, 'I', '1053-1054, 'I', '1055-1056, 'I', '1057-1058, 'I', '1059-1060, 'I', '1061-1062, 'I', '1063-1064, 'I', '1065-1066, 'I', '1067-1068, 'I', '1069-1070, 'I', '1071-1072, 'I', '1073-1074, 'I', '1075-1076, 'I', '1077-1078, 'I', '1079-1080, 'I', '1081-1082, 'I', '1083-1084, 'I', '1085-1086, 'I', '1087-1088, 'I', '1089-1090, 'I', '1091-1092, 'I', '1093-1094, 'I', '1095-1096, 'I', '1097-1098, 'I', '1099-1100, 'I', '1101-1102, 'I', '1103-1104, 'I', '1105-1106, 'I', '1107-1108, 'I', '1109-1110, 'I', '1111-1112, 'I', '1113-1114, 'I', '1115-1116, 'I', '1117-1118, 'I', '1119-1120, 'I', '1121-1122, 'I', '1123-1124, 'I', '1125-1126, 'I', '1127-1128, 'I', '1129-1130, 'I', '1131-1132, 'I', '1133-1134, 'I', '1135-1136, 'I', '1137-1138, 'I', '1139-1140, 'I', '1141-1142, 'I', '1143-1144, 'I', '1145-1146, 'I', '1147-1148, 'I', '1149-1150, 'I', '1151-1152, 'I', '1153-1154, 'I', '1155-1156, 'I', '1157-1158, 'I', '1159-1160, 'I', '1161-1162, 'I', '1163-1164, 'I', '1165-1166, 'I', '1167-1168, 'I', '1169-1170, 'I', '1171-1172, 'I', '1173-1174, 'I', '1175-1176, 'I', '1177-1178, 'I', '1179-1180, 'I', '1181-1182, 'I', '1183-1184, 'I', '1185-1186, 'I', '1187-1188, 'I', '1189-1190, 'I', '1191-1192, 'I', '1193-1194, 'I', '1195-1196, 'I', '1197-1198, 'I', '1199-1200, 'I', '1201-1202, 'I', '1203-1204, 'I', '1205-1206, 'I', '1207-1208, 'I', '1209-1210, 'I', '1211-1212, 'I', '1213-1214, 'I', '1215-1216, 'I', '1217-1218, 'I', '1219-1220, 'I', '1221-1222, 'I', '1223-1224, 'I', '1225-1226, 'I', '1227-1228, 'I', '1229-1230, 'I', '1231-1232, 'I', '1233-1234, 'I', '1235-1236, 'I', '1237-1238, 'I', '1239-1240, 'I', '1241-1242, 'I', '1243-1244, 'I', '1245-1246, 'I', '1247-1248, 'I', '1249-1250, 'I', '1251-1252, 'I', '1253-1254, 'I', '1255-1256, 'I', '1257-1258, 'I', '1259-1260, 'I', '1261-1262, 'I', '1263-1264, 'I', '1265-1266, 'I', '1267-1268, 'I', '1269-1270, 'I', '1271-1272, 'I', '1273-1274, 'I', '1275-1276, 'I', '1277-1278, 'I', '1279-1280, 'I', '1281-1282, 'I', '1283-1284, 'I', '1285-1286, 'I', '1287-1288, 'I', '1289-1290, 'I', '1291-1292, 'I', '1293-1294, 'I', '1295-1296, 'I', '1297-1298, 'I', '1299-1300, 'I', '1301-1302, 'I', '1303-1304, 'I', '1305-1306, 'I', '1307-1308, 'I', '1309-1310, 'I', '1311-1312, 'I', '1313-1314, 'I', '1315-1316, 'I', '1317-1318, 'I', '1319-1320, 'I', '1321-1322, 'I', '1323-1324, 'I', '1325-1326, 'I', '1327-1328, 'I', '1329-1330, 'I', '1331-1332, 'I', '1333-1334, 'I', '1335-1336, 'I', '1337-1338, 'I', '1339-1340, 'I', '1341-1342, 'I', '1343-1344, 'I', '1345-1346, 'I', '1347-1348, 'I', '1349-1350, 'I', '1351-1352, 'I', '1353-1354, 'I', '1355-1356, 'I', '1357-1358, 'I', '1359-1360, 'I', '1361-1362, 'I', '1363-1364, 'I', '1365-1366, 'I', '1367-1368, 'I', '1369-1370, 'I', '1371-1372, 'I', '1373-1374, 'I', '1375-1376, 'I', '1377-1378, 'I', '1379-1380, 'I', '1381-1382, 'I', '1383-1384, 'I', '1385-1386, 'I', '1387-1388, 'I', '1389-1390, 'I', '1391-1392, 'I', '1393-1394, 'I', '1395-1396, 'I', '1397-1398, 'I', '1399-1400, 'I', '1401-1402, 'I', '1403-1404, 'I', '1405-1406, 'I', '1407-1408, 'I', '1409-1410, 'I', '1411-1412, 'I', '1413-1414, 'I', '1415-1416, 'I', '1417-1418, 'I', '1419-1420, 'I', '1421-1422, 'I', '1423-1424, 'I', '1425-1426, 'I', '1427-1428, 'I', '1429-1430, 'I', '1431-1432, 'I', '1433-1434, 'I', '1435-1436, 'I', '1437-1438, 'I', '1439-1440, 'I', '1441-1442, 'I', '1443-1444, 'I', '1445-1446, 'I', '1447-1448, 'I', '1449-1450, 'I', '1451-1452, 'I', '1453-1454, 'I', '1455-1456, 'I', '1457-1458, 'I', '1459-1460, 'I', '1461-1462, 'I', '1463-1464, 'I', '1465-1466, 'I', '1467-1468, 'I', '1469-1470, 'I', '1471-1472, 'I', '1473-1474, 'I', '1475-1476, 'I', '1477-1478, 'I', '1479-1480, 'I', '1481-1482, 'I', '1483-1484, 'I', '1485-1486, 'I', '1487-1488, 'I', '1489-1490, 'I', '1491-1492, 'I', '1493-1494, 'I', '1495-1496, 'I', '1497-1498, 'I', '1499-1500, 'I', '1501-1502, 'I', '1503-1504, 'I', '1505-1506, 'I', '1507-1508, 'I', '1509-1510, 'I', '1511-1512, 'I', '1513-1514, 'I', '1515-1516, 'I', '1517-1518, 'I', '1519-1520, 'I', '1521-1522, 'I', '1523-1524, 'I', '1525-1526, 'I', '1527-1528, 'I', '1529-1530, 'I', '1531-1532, 'I', '1533-1534, 'I', '1535-1536, 'I', '1537-1538, 'I', '1539-1540, 'I', '1541-1542, 'I', '1543-1544, 'I', '1545-1546, 'I', '1547-1548, 'I', '1549-1550, 'I', '1551-1552, 'I', '1553-1554, 'I', '1555-1556, 'I', '1557-1558, 'I', '1559-1560, 'I', '1561-1562, 'I', '1563-1564, 'I', '1565-1566, 'I', '1567-1568, 'I', '1569-1570, 'I', '1571-1572, 'I', '1573-1574, 'I', '1575-1576, 'I', '1577-1578, 'I', '1579-1580, 'I', '1581-1582, 'I', '1583-1584, 'I', '1585-1586, 'I', '1587-1588, 'I', '1589-1590, 'I', '1591-1592, 'I', '1593-1594, 'I', '1595-1596, 'I', '1597-1598, 'I', '1599-1600, 'I', '1601-1602, 'I', '1603-1604, 'I', '1605-1606, 'I', '1607-1608, 'I', '1609-1610, 'I', '1611-1612, 'I', '1613-1614, 'I', '1615-1616, 'I', '1617-1618, 'I', '1619-1620, 'I', '1621-1622, 'I', '1623-1624, 'I', '1625-1626, 'I', '1627-1628, 'I', '1629-1630, 'I', '1631-1632, 'I', '1633-1634, 'I', '1635-1636, 'I', '1637-1638, 'I', '1639-1640, 'I', '1641-1642, 'I', '1643-1644, 'I', '1645-1646, 'I', '1647-1648, 'I', '1649-1650, 'I', '1651-1652, 'I', '1653-1654, 'I', '1655-1656, 'I', '1657-1658, 'I', '1659-1660, 'I', '1661-1662, 'I', '1663-1664, 'I', '1665-1666, 'I', '1667-1668, 'I', '1669-1670, 'I', '1671-1672, 'I', '1673-1674, 'I', '1675-1676, 'I', '1677-1678, 'I', '1679-1680, 'I', '1681-1682, 'I', '1683-1684, 'I', '1685-1686, 'I', '1687-1688, 'I', '1689-1690, 'I', '1691-1692, 'I', '1693-1694, 'I', '1695-1696, 'I', '1697-1698, 'I', '1699-1700, 'I', '1701-1702, 'I', '1703-1704, 'I', '1705-1706, 'I', '1707-1708, 'I', '1709-1710, 'I', '1711-1712, 'I', '1713-1714, 'I', '1715-1716, 'I', '1717-1718, 'I', '1719-1720, 'I', '1721-1722, 'I', '1723-1724, 'I', '1725-1726, 'I', '1727-1728, 'I', '1729-1730, 'I', '1731-1732, 'I', '1733-1734, 'I', '1735-1736, 'I', '1737-1738, 'I', '1739-1740, 'I', '1741-1742, 'I', '1743-1744, 'I', '1745-1746, 'I', '1747-1748, 'I', '1749-1750, 'I', '1751-1752, 'I', '1753-1754, 'I', '1755-1756, 'I', '1757-1758, 'I', '1759-1760, 'I', '1761-1762, 'I', '1763-1764, 'I', '1765-1766, 'I', '1767-1768, 'I', '1769-1770, 'I', '1771-1772, 'I', '1773-1774, 'I', '1775-1776, 'I', '1777-1778, 'I', '1779-1780, 'I', '1781-1782, 'I', '1783-1784, 'I', '1785-1786, 'I', '1787-1788, 'I', '1789-1790, 'I', '1791-1792, 'I', '1793-1794, 'I', '1795-1796, 'I', '1797-1798, 'I', '1799-1800, 'I', '1801-1802, 'I', '1803-1804, 'I', '1805-1806, 'I', '1807-1808, 'I', '1809-1810, 'I', '1811-1812, 'I', '1813-1814, 'I', '1815-1816, 'I', '1817-1818, 'I', '1819-1820, 'I', '1821-1822, 'I', '1823-1824, 'I', '1825-1826, 'I', '1827-1828, 'I', '1829-1830, 'I', '1831-1832, 'I', '1833-1834, 'I', '1835-1836, 'I', '1837-1838, 'I', '1839-1840, 'I', '1841-1842, 'I', '1843-1844, 'I', '1845-1846, 'I', '1847-1848, 'I', '1849-1850, 'I', '1851-1852, 'I', '1853-1854, 'I', '1855-1856, 'I', '1857-1858, 'I', '1859-1860, 'I', '1861-1862, 'I', '1863-1864, 'I', '1865-1866, 'I', '1867-1868, 'I', '1869-1870, 'I', '1871-1872, 'I', '1873-1874, 'I', '1875-1876, 'I', '1877-1878, 'I', '1879-1880, 'I', '1881-1882, 'I', '1883-1884, 'I', '1885-1886, 'I', '1887-1888, 'I', '1889-1890, 'I', '1891-1892, 'I', '1893-1894, 'I', '1895-1896, 'I', '1897-1898, 'I', '1899-1900, 'I', '1901-1902, 'I', '1903-1904, 'I', '1905-1906, 'I', '1907-1908, 'I', '1909-1910, 'I', '1911-1912, 'I', '1913-1914, 'I', '1915-1916, 'I', '1917-1918, 'I', '1919-1920, 'I', '1921-1922, 'I', '1923-1924, 'I', '1925-1926, 'I', '1927-1928, 'I', '1929-1930, 'I', '1931-1932, 'I', '1933-1934, 'I', '1935-1936, 'I', '1937-1938, 'I', '1939-1940, 'I', '1941-1942, 'I', '1943-1944, 'I', '1945-1946, 'I', '1947-1948, 'I', '1949-1950, 'I', '1951-1952, 'I', '1953-1954, 'I', '1955-1956, 'I', '1957-1958, 'I', '1959-1960, 'I', '1961-1962, 'I', '1963-1964, 'I', '1965-1966, 'I', '1967-1968, 'I', '1969-1970, 'I', '1971-1972, 'I', '1973-1974, 'I', '1975-1976, 'I', '1977-1978, 'I', '1979-1980, 'I', '1981-1982, 'I', '1983-1984, 'I', '1985-1986, 'I', '1987-1988, 'I', '1989-1990, 'I', '1991-1992, 'I', '1993-1994, 'I', '1995-1996, 'I', '1997-1998, 'I', '1999-2000, 'I', '2001-2002, 'I', '2003-2004, 'I', '2005-2006, 'I', '2007-2008, 'I', '2009-2010, 'I', '2011-2012, 'I', '2013-2014, 'I', '2015-2016, 'I', '2017-2018, 'I', '2019-2020, 'I', '2021-2022, 'I', '2023-2024, 'I', '2025-2026, 'I', '2027-2028, 'I', '2029-2030, 'I', '2031-2032, 'I', '2033-2034, 'I', '2035-2036, 'I', '2037-2038, 'I', '2039-2040, 'I', '2041-2042, 'I', '2043-2044, 'I', '2045-2046, 'I', '2047-2048, 'I', '2049-2050, 'I', '2051-2052, 'I', '2053-2054, 'I', '2055-2056, 'I', '2057-2058, 'I', '2059-2060, 'I', '2061-2062, 'I', '2063-2064, 'I', '2065-2066, 'I', '2067-2068, 'I', '2069-2070, 'I', '2071-2072, 'I', '2073-2074, 'I', '2075-2076, 'I', '2077-2078, 'I', '2079-2080, 'I', '2081-2082, 'I', '2083-2084, 'I', '2085-2086, 'I', '2087-2088, 'I', '2089-2090, 'I', '2091-2092, 'I', '2093-2094, 'I', '2095-2096, 'I', '2097-2098, 'I', '2099-2100, 'I', '2101-2102, 'I', '2103-2104, 'I', '2105-2106, 'I', '2107-2108, 'I', '2109-2110, 'I', '2111-2112, 'I', '2113-2114, 'I', '2115-2116, 'I', '2117-2118, 'I', '2119-2120, 'I', '2121-2122, 'I', '2123-2124, 'I', '2125-2126, 'I', '2127-2128, 'I', '2129-2130, 'I', '2131-2132, 'I', '2133-2134, 'I', '2135-2136, 'I', '2137-2138, 'I', '2139-2140, 'I', '2141-2142, 'I', '2143-2144, 'I', '2145-2146, 'I', '2147-2148, 'I', '2149-2150, 'I', '2151-2152, 'I', '2153-2154, 'I', '2155-2156, 'I', '2157-2158, 'I', '2159-2160, 'I', '2161-2162, 'I', '2163-2164, 'I', '2165-2166, 'I', '2167-2168, 'I', '2169-2170, 'I', '2171-2172, 'I', '2173-2174, 'I', '2175-2176, 'I', '2177-2178, 'I', '2179-2180, 'I', '2181-2182, 'I', '2183-2184, 'I', '2185-2186, 'I', '2187-2188, 'I', '2189-2190, 'I', '2191-2192, 'I', '2193-2194, 'I', '2195-2196, 'I', '2197-2198, 'I', '2199-2200, 'I', '2201-2202, 'I', '2203-2204, 'I', '2205-2206, 'I', '2207-2208, 'I', '2209-2210, 'I', '2211-2212, 'I', '2213-2214, 'I', '2215-2216, 'I', '2217-2218, 'I', '2219-2220, 'I', '2221-2222, 'I', '2223-2224, 'I', '2225-2226, 'I', '2227-2228, 'I', '2229-2230, 'I', '2231-2232, 'I', '2233-2234, 'I', '2235-2236, 'I', '2237-2238, 'I', '2239-2240, 'I', '2241-2242, 'I', '2243-2244, 'I', '2245-2246, 'I', '2247-2248, 'I', '2249-2250, 'I', '2251-2252, 'I', '2253-2254, 'I', '2255-2256, 'I', '2257-2258, 'I', '2259-2260, 'I', '2261-2262, 'I', '2263-2264, 'I', '2265-2266, 'I', '2267-2268, 'I', '2269-2270, 'I', '2271-2272, 'I', '2273-2274, 'I', '2275-2276, 'I', '2277-2278, 'I', '2279-2280, 'I', '2281-2282, 'I', '2283-2284, 'I', '2285-2286, 'I', '2287-2288, 'I', '2289-2290, 'I', '2291-2292, 'I', '2293-2294, 'I', '2295-2296, 'I', '2297-2298, 'I', '2299-2300, 'I', '2301-2302, 'I', '2303-2304, 'I', '2305-2306, 'I', '2307-2308, 'I', '2309-2310, 'I', '2311-2312, 'I', '2313-2314, 'I', '2315-2316, 'I', '2317-2318, 'I', '2319-2320, 'I', '2321-2322, 'I', '2323-2324, 'I', '2325-2326, 'I', '2327-2328, 'I', '2329-2330, 'I', '2331-2332, 'I', '2333-2334, 'I', '2335-2336, 'I', '2337-2338, 'I', '2339-2340, 'I', '2341-2342, 'I', '2343-2344, 'I', '2345-2346, 'I', '2347-2348, 'I', '2349-2350, 'I', '2351-2352, 'I', '2353-2354, 'I', '2355-2356, 'I', '2357-2358, 'I', '2359-2360, 'I', '2361-2362, 'I', '2363-2364, 'I', '2365-2366, 'I', '2367-2368, 'I', '2369-2370, 'I', '2371-2372, 'I', '2373-2374, 'I', '2375-2376, 'I', '2377-2378, 'I', '2379-2380, 'I', '2381-2382, 'I', '2383-2384, 'I', '2385-2386, 'I', '2387-2388, 'I', '2389-2390, 'I', '2391-2392, 'I', '2393-2394, 'I', '2395-2396, 'I', '2397-2398, 'I', '2399-2400, 'I', '2401-2402, 'I', '2403-2404, 'I', '2405-2406, 'I', '2407-2408, 'I', '2409-2410, 'I', '2411-2412, 'I', '2413-2414, 'I', '2415-2416, 'I', '2417-2418, 'I', '2419-2420, 'I', '2421-2422, 'I', '2423-2424, 'I', '2425-2426, 'I', '2427-2428, 'I', '2429-2430, 'I', '2431-2432, 'I', '2433-2434, 'I', '2435-2436, 'I', '2437-2438, 'I', '2439-2440, 'I', '2441-2442, 'I', '2443-2444, 'I', '2445-2446, 'I', '2447-2448, 'I', '2449-2450, 'I', '2451-2452, 'I', '2453-2454, 'I', '2455-2456, 'I', '2457-2458, 'I', '2459-2460, 'I', '2461-2462, 'I', '2463-2464, 'I', '2465-2466, 'I', '2467-2468, 'I', '2469-2470, 'I', '2471-2472, 'I', '2473-2474, 'I', '2475-2476, 'I', '2477-2478, 'I', '2479-2480, 'I', '2481-2482, 'I', '2483-2484, 'I', '2485-2486, 'I', '2487-2488, '



Best Local Similarity 16.7%; Pred. No. 2.1e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 192 CTALAGSSSFTC 203

## RESULT 42

T31345

hypothetical protein G01D9.5 - Caenorhabditis briggsae

C/Species: Caenorhabditis briggsae

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T31345

R:Waterston, R.

submitted to the EMBL Data Library, April 1996

A/Description: The C. briggsae genome sequencing project.

A/Reference number: Z21010

A/Accession: T31345

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-4767 &lt;WAT&gt;

A/Cross-references: UNIPROT:Q17301; UNIPARC:UPI000011013B; EMBL:U56248; NID:G1233789; P

C/Genetics:

A/Introns: 269/3; 341/3; 853/1; 920/2; 4452/3; 4534/3; 4592/3; 4654/3; 4670/3; 4707/2; 4

A/Note: G01D9.5

C/Keywords: carrier protein; phosphopantetheine; phosphoprotein

F/3472-3541/Domain: acyl carrier protein homology &lt;ACP1&gt;

F/4039-4427/Domain: acetate-CoA ligase homology &lt;ACL&gt;

F/4447-4514/Domain: acyl carrier protein homology &lt;ACP2&gt;

F/2210,3505/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 53.6%; Score 15; DB 2; Length 4767;  
Best Local Similarity 16.7%; Pred. No. 2.4e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 1423 CSSSFHALSSAC 1434

## RESULT 43

T15789

hypothetical protein C41A3.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C/Accession: T15789

R:Bentley, D.

submitted to the EMBL Data Library, November 1995

A/Description: The sequence of C. elegans cosmid C41A3.

A/Reference number: Z18404

A/Accession: T15789

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-7829 &lt;BEN&gt;

A/Cross-references: UNIPROT:Q18559; UNIPARC:UPI000011014C; EMBL:U41541; NID:G1109867; P

C/Genetics:

A/Gene: CBSP\_C41A3.1

A/Introns: 29/2; 69/3; 141/2; 192/2; 271/3; 321/2; 1000/1; 1098/1; 1127/3; 1168/2; 1282/

/2; 7514/3; 7596/3; 7654/3; 7716/3; 7732/3; 7769/2; 7797/3

C/Keywords: carrier protein; phosphopantetheine; phosphoprotein

F/6547-6616/Domain: acyl carrier protein homology &lt;ACP&gt;

F/2833,5521,6580/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 53.6%; Score 15; DB 2; Length 7829;  
Best Local Similarity 16.7%; Pred. No. 2.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 4475 CSSSFYALSSAC 4486

RESULT 44  
I38346  
elastic titin - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jul-2004

C/Accession: I38346

R:Labelt, S.; Kolmerer, B.

Science 270, 293-296, 1995

A/Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.

A/Reference number: A57430; MUID:96026330; PMID:7569978

A/Accession: I38346

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-7962 &lt;RES&gt;

A/Cross-references: UNIPROT:Q10465; UNIPARC:UPI000011010C; EMBL:X90569; NID:G1017426; P

C/Genetics:

A/Gene: GDB:TTN

A/Cross-references: GDB:127867; OMIM:188840

A/Map position: 2q31-2q31

Query Match 53.6%; Score 15; DB 2; Length 7962;  
Best Local Similarity 16.7%; Pred. No. 2.6e-05;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 381 CTASNEAGSSSC 392

## RESULT 45

R-phycoerythrin beta-2 chain - red alga (Gastrocionium coulteri) (fragment)

C/Species: Gastrocionium coulteri

C/Date: 07-Mar-1988 #sequence\_revision 07-Mar-1988 #text\_change 09-Jul-2004

C/Accession: D22565

R:Klotz, A.V.; Glazer, A.N.

J. Biol. Chem. 260, 4856-4863, 1985

A/Title: Characterization of the bilin attachment sites in R-phycoerythrin.

A/Reference number: A22565; MUID:85182601; PMID:3886644

A/Accession: D22565

A/Molecule type: protein

A/Residues: 1-21 &lt;KLO&gt;

A/Cross-references: UNIPROT:Q7M276; UNIPARC:UPI0000174E0D

C/Superfamily: phycocyanin

Query Match 50.0%; Score 14; DB 2; Length 21;  
Best Local Similarity 16.7%; Pred. No. 0.0004;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

Db 8 CMLSSVASYC 19

## RESULT 46

G82754

hypothetical protein X90861 [imported] - Xylella fastidiosa (strain 9a5c)

C/Species: Xylella fastidiosa

C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C/Accession: G82754

C/Anonymous: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A/Reference number: A82515; MUID:20365717; PMID:10910347

A/Note: For a complete list of authors see reference number A59328 below

A/Accession: G82754

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-88 &lt;SIM&gt;

A/Cross-references: UNIPROT:Q9PF17; UNIPARC:UPI00000C254B; GB:AE003925; GB:AE003849; NIT

A/Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;

Briano, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carter, I

as-Neto, E.; Docena, C.; El-Dorfy, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Klitajma, J.P.; Krieger, J.E.; Kuramae, E.E.; Laizy  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.  
A:Authors: Martins, E.M.F.; Matsushima, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.  
Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF0861

Query Match 50.0%; Score 14; DB 2; Length 88;  
Best Local Similarity 16.7%; Pred. No. 0.00052;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 23 CSSSVTLKSSRC 34

RESULT 47  
154781  
fibroblast growth factor receptor FGFR-1, secreted splice form - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: T54781  
R:Leung, H.Y.; Gullick, W.J.; Lemoine, N.R.  
Int. J. Cancer 59, 667-675, 1994  
A:Title: Expression and functional activity of fibroblast growth factors and their recep  
A:Reference number: 154781; MUID:95048906; PMID:7960240  
A:Accession: T54781  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-92 <RES>  
A:Cross-references: UNIPARC:UPI0000116DB9; GB:S76658; NID:9913557; PIDN:AAD14230.1; PID:  
A:Experimental source: pancreatic cancer cell line PT45  
C:Genetics:

A:Gene: GDB:RGFR1; FLT2  
A:Cross-references: GDB:119913; OMIM:136350  
A:Map position: 8p11.2-8p11.1  
A:Introns: 58/3  
C:Keywords: growth factor receptor

Query Match 50.0%; Score 14; DB 2; Length 92;  
Best Local Similarity 16.7%; Pred. No. 0.00053;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 30 CGAGSASSSLC 41

RESULT 48  
T28947  
hypothetical protein F07C4.11 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T28947  
R:Miller, N.; Steillyes, L.  
submitted to the EMBL Data Library, January 1997  
A:Description: The sequence of C. elegans cosmid F07C4.  
A:Reference number: Z20546  
A:Accession: T28947  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-97 <ML>  
A:Cross-references: UNIPROT:P91216; UNIPARC:UPI00000737D; EMBL:U80023; PIDN:AAC48019.1;  
A:Experimental source: strain Bristol N2; clone F07C4  
C:Genetics:

A:Gene: CESP:F07C4.11  
A:Map position: 5

Query Match 50.0%; Score 14; DB 2; Length 97;  
Best Local Similarity 16.7%; Pred. No. 0.00053;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 56 CTTVAADSSPSC 67

RESULT 49  
C72775  
hypothetical protein APE0189 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: C72775  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah  
awa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: C72775  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-100 <KAW>  
A:Cross-references: UNIPROT:Q9YFR0; UNIPARC:UPI000005DA3D; DBJ:AP000058; NID:95103388;  
A:Experimental source: strain K1  
C:Genetics:

A:Gene: APE0189  
C:Superfamily: Aeropyrum pernix hypothetical protein APE0189

Query Match 50.0%; Score 14; DB 2; Length 100;  
Best Local Similarity 16.7%; Pred. No. 0.00054;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 65 CASVSRMRASSC 76

RESULT 50  
T14780  
hypothetical protein DKFZ566B1346.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T14780  
R:Oltmannselder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, August 1999  
A:Reference number: Z18184  
A:Accession: T14780  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-101 <OTT>  
A:Cross-references: UNIPARC:UPI000016AC6B; EMBL:AL110253  
A:Experimental source: fetal kidney; clone DKFZ566B1346  
C:Genetics:

A:Note: DKFZ566B1346.1

Query Match 50.0%; Score 14; DB 2; Length 101;  
Best Local Similarity 16.7%; Pred. No. 0.00054;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 38 CSISVVTQTASC 49

Search completed: January 4, 2006, 16:10:30  
Job time : 13.4174 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 4, 2006, 15:02:48 ; Search time 78.3652 Seconds  
(without alignments)  
108.037 Million cell updates/sec

Title: US-09-932-322-12  
Perfect score: 28  
Sequence: 1 CXXXXXXXXXXC 12

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%  
Listing first 1000 summaries

Database : UniProt 05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18	64.3	307	1	SGS3 DROME
2	18	64.3	347	2	Q8K340 MOUSE
3	18	64.3	362	2	Q8CDD3 MOUSE
4	18	64.3	367	2	Q6AY38 ORYSA
5	18	64.3	384	2	Q8BL43 MOUSE
6	18	64.3	730	2	Q8LPI3 PEA
7	17	60.7	35	2	Q9LQ12 ARATH
8	17	60.7	137	2	Q24960 GIALLA
9	17	60.7	137	2	Q24987 GIALLA
10	17	60.7	210	2	Q7XHS1 ORYSA
11	17	60.7	234	2	Q9XY90 GIALLA
12	17	60.7	245	2	Q18039 CAELH
13	17	60.7	306	2	Q5MQW8 CRYNE
14	17	60.7	378	2	Q6ASQ2 ORYSA
15	17	60.7	396	2	Q4LGT2 9BURY
16	17	60.7	429	2	Q8CXI1 MOUSE
17	17	60.7	540	2	Q93H33 STRAW
18	17	60.7	667	2	Q9XTK3 GIALLA
19	17	60.7	1031	2	Q5B3H3 EMBENI
20	17	60.7	5703	1	MUC5B HUMAN
21	16	57.1	93	2	Q6EUA6 ORYSA
22	16	57.1	100	2	Q850L9 ORYSA
23	16	57.1	100	2	Q62YJ1 BURWA
24	16	57.1	110	1	CBP1 ABECA
25	16	57.1	144	2	Q5C209 SCBPA
26	16	57.1	152	2	Q9YBV0 ABRPE
27	16	57.1	161	2	Q7SLM9 ORYSA
28	16	57.1	172	2	Q7PY44 ANOGA
29	16	57.1	176	2	Q6ZD90 ORYSA
30	16	57.1	187	2	Q55014 DICDI
31	16	57.1	208	2	Q82LKO STRAW

32	16	57.1	210	2	Q67V9 ORYSA
33	16	57.1	215	2	Q5K9J0 CRYNE
34	16	57.1	214	2	Q55K09 CRYNE
35	16	57.1	215	2	Q5K919 CRYNE
36	16	57.1	227	2	Q4OSJ9 9LECA
37	16	57.1	230	2	Q70NB5 GIALLA
38	16	57.1	234	2	Q8LD53 ARATH
39	16	57.1	234	2	Q9F37 ARATH
40	16	57.1	244	2	Q16421 CAELH
41	16	57.1	247	2	Q66B6 CAELH
42	16	57.1	248	2	Q16424 CAELH
43	16	57.1	257	2	Q9XV90 CAELH
44	16	57.1	262	2	Q8CAZ5 MOUSE
45	16	57.1	269	2	Q7PRQ7 ANOGA
46	16	57.1	278	2	Q6ZUB7 HUMAN
47	16	57.1	285	2	Q8PFE6 XANAC
48	16	57.1	286	2	Q6ZQO1 HUMAN
49	16	57.1	287	2	Q9F934 BIPAD
50	16	57.1	297	2	Q89DY5 BRAJA
51	16	57.1	305	2	Q5OLO5 ORYSA
52	16	57.1	309	2	Q5ZBR6 ORYSA
53	16	57.1	320	2	Q8N4O3 HUMAN
54	16	57.1	325	2	Q91E43 9ALPH
55	16	57.1	329	2	Q7WBR6 BORPA
56	16	57.1	329	2	Q7WPO8 BORBR
57	16	57.1	353	2	Q8XQX5 RALSO
58	16	57.1	366	2	Q55057 CRYNE
59	16	57.1	381	2	Q5ZPV8 ORYSA
60	16	57.1	381	2	Q71021 ORYSA
61	16	57.1	426	2	Q941V9 ORYSA
62	16	57.1	437	2	Q7XUJ0 ORYSA
63	16	57.1	448	2	Q8BP04 MOUSE
64	16	57.1	465	2	Q7SLH3 NEUCR
65	16	57.1	471	2	Q4PD71 USYMA
66	16	57.1	499	2	Q6S60 ORYSA
67	16	57.1	504	2	Q9XZ8 LEIMA
68	16	57.1	534	2	Q4LZH5 9BURY
69	16	57.1	538	2	Q86XHO HUMAN
70	16	57.1	539	1	GSHB ARATH
71	16	57.1	543	2	Q4QEB1 LEIMA
72	16	57.1	544	2	Q9F353 STRCO
73	16	57.1	556	2	Q7SN76 ACATE
74	16	57.1	569	2	Q7GXT3 GIALLA
75	16	57.1	622	2	Q5JVP6 HUMAN
76	16	57.1	667	1	TS11 GIALLA
77	16	57.1	709	2	Q9744 GIALLA
78	16	57.1	717	2	Q4SB79 TETNG
79	16	57.1	719	2	Q9U019 GIALLA
80	16	57.1	729	2	Q9KH34 9GAMM
81	16	57.1	760	2	Q7OU65 GIALLA
82	16	57.1	772	2	Q5XXR3 RAT
83	16	57.1	776	1	ARHG6 HUMAN
84	16	57.1	776	2	Q7Z3W1 HUMAN
85	16	57.1	776	2	Q5Y066 HUMAN
86	16	57.1	799	2	Q4T7H8 TETNG
87	16	57.1	87	2	Q7XDB6 ORYSA
88	16	57.1	868	2	CHS1 CRYNY
89	16	57.1	1041	1	Q4SDRO TETNG
90	16	57.1	1046	2	Q4QBW2 LEIMA
91	16	57.1	1107	2	Q5CV70 CRYPV
92	16	57.1	1107	2	Q6CBRS YARLI
93	16	57.1	1136	2	Q57ZRO 9RYP
94	16	57.1	1167	2	Q6MM93 ORYSA
95	16	57.1	1178	2	Q5QDB2 CRYNE
96	16	57.1	1236	2	Q4HV91 GIBZE
97	16	57.1	1341	2	Q4RL73 TETNG
98	16	57.1	1494	2	Q4P4P1 USYMA
99	16	57.1	1519	2	Q4Q998 LEIMA
100	16	57.1	1509	2	Q6ID66 TOXCO
101	16	57.1	5072	2	Q5YUT7 9COCO
102	16	53.6	35	2	Q5VUT3 9COCO
103	16	53.6	36	2	Q53N14 ORYSA
104	15	53.6	43	2	

Q67V9 ORYSA	Q5K9J0 CRYNE	Q55K09 CRYNE	Q5K919 CRYNE	Q4OSJ9 9LECA	Q70NB5 GIALLA	Q8LD53 ARATH	Q9F37 ARATH	Q16421 CAELH	Q66B6 CAELH	Q16424 CAELH	Q9XV90 CAELH	Q8CAZ5 MOUSE	Q7PRQ7 ANOGA	Q6ZUB7 HUMAN	Q8PFE6 XANAC	Q6ZQO1 HUMAN	Q9F934 BIPAD	Q89DY5 BRAJA	Q5OLO5 ORYSA	Q5ZBR6 ORYSA	Q8N4O3 HUMAN	Q91E43 9ALPH	Q7WBR6 BORPA	Q7WPO8 BORBR	Q8XQX5 RALSO	Q55057 CRYNE	Q5ZPV8 ORYSA	Q71021 ORYSA	Q941V9 ORYSA	Q7XUJ0 ORYSA	Q8BP04 MOUSE	Q7SLH3 NEUCR	Q4PD71 USYMA	Q6S60 ORYSA	Q9XZ8 LEIMA	Q4LZH5 9BURY	Q86XHO HUMAN	GSHB ARATH	Q4QEB1 LEIMA	Q9F353 STRCO	Q7SN76 ACATE	Q7GXT3 GIALLA	Q5JVP6 HUMAN	TS11 GIALLA	Q9744 GIALLA	Q4SB79 TETNG	Q9U019 GIALLA	Q9KH34 9GAMM	Q7OU65 GIALLA	Q5XXR3 RAT	ARHG6 HUMAN	Q7Z3W1 HUMAN	Q5Y066 HUMAN	Q4T7H8 TETNG	Q7XDB6 ORYSA	CHS1 CRYNY	Q4SDRO TETNG	Q4QBW2 LEIMA	Q5CV70 CRYPV	Q6CBRS YARLI	Q57ZRO 9RYP	Q6MM93 ORYSA	Q5QDB2 CRYNE	Q4HV91 GIBZE	Q4RL73 TETNG	Q4P4P1 USYMA	Q4Q998 LEIMA	Q6ID66 TOXCO	Q5YUT7 9COCO	Q5VUT3 9COCO	Q53N14 ORYSA
-------------	--------------	--------------	--------------	--------------	---------------	--------------	-------------	--------------	-------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	-------------	-------------	--------------	--------------	------------	--------------	--------------	--------------	---------------	--------------	-------------	--------------	--------------	---------------	--------------	---------------	------------	-------------	--------------	--------------	--------------	--------------	------------	--------------	--------------	--------------	--------------	-------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------

105	15	53.6	57	2	094MN5_9CAUD	094MN5_bacterioph	178	15	53.6	227	2	09VA83_DROME	09VA83_drosophila
106	15	53.6	57	2	04MOU9_BACCE	04mu9 bacillus ce	179	15	53.6	229	2	072969_9PUNG	072969_fungal endo
107	15	53.6	57	2	04MOV1_BACCE	04mov1 bacillus ce	180	15	53.6	229	2	072973_9PUNG	072973_fungal endo
108	15	53.6	72	2	05C4E3_SCHJA	05c4e3 schistosoma	181	15	53.6	229	2	0729K3_9PUNG	0729K3_fungal endo
109	15	53.6	72	2	04MOV6_BACCE	04mov6 bacillus ce	182	15	53.6	229	2	0729K4_9PUNG	0729K4_fungal endo
110	15	53.6	86	2	09YV33_9HIV1	09yv33 human immun	183	15	53.6	230	2	056P86_9METZ	056P86_uncultured
111	15	53.6	87	2	07S7M4_NEUCR	07s7m4 neurospora	184	15	53.6	234	2	056P87_9PUNG	056P87_fungal sp.
112	15	53.6	87	2	04RQO2_TERNG	04rq2 tetraodon n	185	15	53.6	236	2	056P94_9METZ	056P94_uncultured
113	15	53.6	94	2	06IG47_DROME	06ig47 drosophila	186	15	53.6	239	2	096UV6_ORPRL	096UV6_optioetoma
114	15	53.6	94	2	09DKL7_9VIRU	09dkl7 spodiopetra	187	15	53.6	242	2	051T33_MAGGR	051T33_magnaporthe
115	15	53.6	98	2	07GPF3_GIALA	07gpf3 giardia lam	188	15	53.6	243	2	065XAS_ORYSA	065XAS_oryza sativ
116	15	53.6	99	2	06UTV5_CRYPE	06utv5 crinipellis	189	15	53.6	245	2	08FBS9_ECOLB	08FBS9_escherichia
117	15	53.6	99	2	04SDO4_TERNG	04sdo4 tetraodon n	190	15	53.6	248	2	061HW9_CAEBR	061HW9_caenorhabdi
118	15	53.6	106	2	06TSJ1_PABBR	06tsj1 paracoccidi	191	15	53.6	248	2	045450_CAEEL	045450_caenorhabdi
119	15	53.6	107	2	06HEN9_ORYSA	06hen9 oryza sativ	192	15	53.6	249	2	0624V6_CAEBR	0624V6_caenorhabdi
120	15	53.6	117	2	06HSN3_ORYSA	06hsn3 oryza sativ	193	15	53.6	250	2	0624V4_CAEBR	0624V4_caenorhabdi
121	15	53.6	125	2	084N44_9FABA	084n44 dialium gui	194	15	53.6	252	2	061135_CAEBR	061135_caenorhabdi
122	15	53.6	128	2	04LI18_9BURK	04li18 burkholderi	195	15	53.6	252	2	0624V5_CAEBR	0624V5_caenorhabdi
123	15	53.6	132	2	08R5F9_MOUSE	08r5f9 mus musculu	196	15	53.6	257	2	07P2D8_ANGCA	07P2D8_anopheles g
124	15	53.6	139	2	072A85_DESVH	072a85 desulfovibr	197	15	53.6	257	2	0676U4_9SOLA	0676U4_nicotiana a
125	15	53.6	141	2	0825O9_STRAM	0825o9 streptomyce	198	15	53.6	261	2	06FPP4_CANGA	06FPP4_candida gla
126	15	53.6	150	2	05Z6U8_ORYSA	05z6u8 oryza sativ	199	15	53.6	264	2	04Q805_LEIMA	04Q805_leishmania
127	15	53.6	152	2	06R217_9PEZI	06r217 arthrinitum	200	15	53.6	264	2	05DQ89_9CORA	05DQ89_canine coro
128	15	53.6	153	2	063176_BURPS	063176 burkholderi	201	15	53.6	268	2	022798_ARATH	022798_arabidopsis
129	15	53.6	154	2	066R16_YERPE	066r16 yerinia ps	202	15	53.6	269	2	06ZFE5_ORYSA	06ZFE5_oryza sativ
130	15	53.6	154	2	08ZB15_YERPE	08zb15 yerinia pe	203	15	53.6	273	2	041070_PEA	041070_pisum sativ
131	15	53.6	155	2	07YSR5_DROME	07ysr5 drosophila	204	15	53.6	281	2	05RX10_ANGCA	05RX10_anopheles g
132	15	53.6	155	2	06Y1L9_DROYA	06y1l9 drosophila	205	15	53.6	281	2	07RKP8_PLAYO	07RKP8_plasmodium
133	15	53.6	155	2	0721J0_DROSI	0721j0 drosophila	206	15	53.6	282	2	04LH85_9BURK	04LH85_burkholderi
134	15	53.6	155	2	0721J1_DROME	0721j1 drosophila	207	15	53.6	282	2	08UFM9_AGRYS	08UFM9_agrobacteri
135	15	53.6	155	2	07XEA0_ORYSA	07xea0 oryza sativ	208	15	53.6	285	2	05LXE3_SILPO	05LXE3_silicibacte
136	15	53.6	155	2	093TQ7_9DELT	093tq7 uncultured	209	15	53.6	287	2	043528_LYCES	043528_lycopersico
137	15	53.6	156	2	059LP3_CANAL	059lp3 candida alb	210	15	53.6	289	2	043527_LYCES	043527_lycopersico
138	15	53.6	163	2	09B1M0_GIALA	09b1m0 giardia lam	211	15	53.6	293	2	061YT3_CAEBR	061YT3_caenorhabdi
139	15	53.6	163	2	07MQD6_VIBVY	07mqd6 vibrio vuln	212	15	53.6	294	2	09VTR3_DROME	09VTR3_drosophila
140	15	53.6	165	2	06R216_9PEZI	06r216 apiopora m	213	15	53.6	306	2	07F8S0_ORYSA	07F8S0_oryza sativ
141	15	53.6	165	2	08DD94_VIBVU	08dd94 vibrio vuln	214	15	53.6	318	2	07CZES_AGRYS	07CZES_agrobacteri
142	15	53.6	169	2	04MHT3_BACCE	04mht3 bacillus ce	215	15	53.6	332	2	062009_CERCA	062009_ceratitis c
143	15	53.6	172	2	04NFB6_9MICA	04nfb6 arthropacte	216	15	53.6	334	2	08S202_DROME	08S202_drosophila
144	15	53.6	177	2	07UGY7_RHODA	07ugy7 rhodopirell	217	15	53.6	334	2	052V10_CIOIN	052V10_ciona intes
145	15	53.6	179	2	05Z2L9_MYCHI	05z2l9 mycoplasma	218	15	53.6	335	2	09VW95_DROME	09VW95_drosophila
146	15	53.6	181	2	061UD0_DROME	061ud0 anaeromyxob	219	15	53.6	335	2	052VH9_CIOIN	052VH9_ciona intes
147	15	53.6	183	2	0833W8_HORYD	0833w8 drosophila	220	15	53.6	341	2	052V12_CIOIN	052V12_ciona intes
148	15	53.6	189	2	082P67_STRAM	082p67 streptomyce	221	15	53.6	341	2	056N31_CIOIN	056N31_ciona intes
149	15	53.6	192	2	05KRJ2_CRYNE	05krj2 cryptococcu	222	15	53.6	342	2	053PA6_ORYSA	053PA6_oryza sativ
150	15	53.6	192	2	05SG00_CRYNE	05sg00 cryptococcu	223	15	53.6	343	2	091FHE_9INFA	091FHE_influenza a
151	15	53.6	193	2	07XH6_ORYSA	07xh6 oryza sativ	224	15	53.6	343	2	04YD4_TERNG	04YD4_tetradon n
152	15	53.6	193	2	07OOG1_GIALA	07oog1 giardia lam	225	15	53.6	347	2	08SYW5_DROME	08SYW5_drosophila
153	15	53.6	195	2	053M55_ORYSA	053m55 oryza sativ	226	15	53.6	347	2	08G4A1_BIFLO	08G4A1_bifidobacte
154	15	53.6	196	2	061G98_DROME	061g98 drosophila	227	15	53.6	350	2	0467B8_TERNG	0467B8_tetradon n
155	15	53.6	199	2	04OG66_9LECA	04og66 pertusaria	228	15	53.6	352	2	06P898_XENTR	06P898_xenopus tro
156	15	53.6	209	2	06V282_9BACT	06v282 symbiont ba	229	15	53.6	354	2	061LP0_DROME	061LP0_drosophila
157	15	53.6	213	2	06OQR7_CAEBR	06oqr7 caenorhabdi	230	15	53.6	357	2	09YV87_CAEEL	09YV87_caenorhabdi
158	15	53.6	214	2	04OS18_9LECA	04os18 pertusaria	231	15	53.6	357	2	MESD1_HUMAN	09h1k6 homo sapien
159	15	53.6	215	2	07OY29_GIALA	07oy29 giardia lam	232	15	53.6	362	1	MESD1_MOUSE	09h1k6 homo sapien
160	15	53.6	217	2	07OY29_GIALA	07oy29 giardia lam	233	15	53.6	362	1	07OJL0_ANGCA	07OJL0_anopheles g
161	15	53.6	218	1	CK5F2_RAT	07j1h5 rattus norv	234	15	53.6	362	2	04V3U8_DROME	04V3U8_drosophila
162	15	53.6	220	2	071T10_CHICK	071t10 gallus gall	235	15	53.6	362	2	05B2V3_RAT	05B2V3_rattus norv
163	15	53.6	224	2	05KCO4_CRYNE	05kco4 cryptococcu	236	15	53.6	362	2	0542D0_MOUSE	0542D0_mus musculu
164	15	53.6	225	2	04QSH8_9LECA	04qsh8 pertusaria	237	15	53.6	362	2	08CFLO_MOUSE	08CFLO_mus musculu
165	15	53.6	225	2	055XU2_CRYNE	055xu2 cryptococcu	238	15	53.6	371	2	0510P9_HUMAN	0510P9_homo sapien
166	15	53.6	226	1	C553_PARDE	P29967 paracoccus	239	15	53.6	372	2	09X1Y6_ORYSA	09X1Y6_oryza sativ
167	15	53.6	227	2	051672_PARDE	051672 paracoccus	240	15	53.6	377	2	07SEPF_ASHGO	07SEPF_ashbya gos
168	15	53.6	227	2	04QSG3_9LECA	04qsg3 pertusaria	241	15	53.6	385	2	07WYF3_TOXGO	07WYF3_toxoplasma
169	15	53.6	227	2	04QGM6_9ASCO	04qgm6 pertusaria	242	15	53.6	386	2	09VEY9_DROME	09VEY9_drosophila
170	15	53.6	227	2	04QSI8_9LECA	04qsi8 pertusaria	243	15	53.6	387	2	08KO10_BURCE	08KO10_burkholderi
171	15	53.6	227	2	04QSG8_9LECA	04qsg8 pertusaria	244	15	53.6	387	2	08KOJ7_BURCE	08KOJ7_burkholderi
172	15	53.6	227	2	04QSR6_9LECA	04qsr6 pertusaria	245	15	53.6	387	2	09B3Y6_ORPPO	09B3Y6_ophiocara p
173	15	53.6	227	2	04QSG0_9LECA	04qsg0 pertusaria	246	15	53.6	398	2	09B3Y2_9GOBI	09B3Y2_acanthogobi
174	15	53.6	227	2	04QSI6_9LECA	04qsi6 pertusaria	247	15	53.6	398	2	09B6C2_LATUJA	09B6C2_latolabdrax
175	15	53.6	227	2	05JUE7_HUMAN	05jue7 homo sapien	248	15	53.6	398	2	09BA00_9GOBI	09BA00_bostrychus
176	15	53.6	227	2	09UT7A3_DROME	09ut7a3 drosophila	249	15	53.6	398	2		
177	15	53.6	227	2			250	15	53.6	398	2		

251	15	53.6	407	2	Q8SG80_MICSA	Q8sg80 micropterus	324	15	53.6	688	2	Q4PGR8_USTWA	Q4pgr8 ustilago ma
252	15	53.6	409	2	Q9SUA2_ARATH	Q9sua2 arabidopsis	325	15	53.6	694	2	Q4T212_TERNG	Q4t212 tetraodon n
253	15	53.6	410	2	Q5HYM7_HUMAN	Q5hyt7 homo sapien	326	15	53.6	695	2	Q7XG03_ORYSA	Q7xg03 oryza sativ
254	15	53.6	411	1	CRPI_YARLI	P45815 yarrowia.l1	327	15	53.6	695	2	Q94HU6_ORYSA	Q94hu6 oryza sativ
255	15	53.6	415	2	Q5TVF2_ANOCA	Q5tvf2 anopheles g	328	15	53.6	705	2	Q7S136_NEUCR	Q7s136 neucpora
256	15	53.6	416	2	Q6RKG7_GIBZE	Q6rkg7 gibberella	329	15	53.6	705	2	Q818W1_GIALA	Q818w1 giardia lam
257	15	53.6	419	1	Q5SRJ3_CRYNE	Q5srj3 cryptococcu	330	15	53.6	713	1	LRPI0_HUMAN	Q7z4f1 homo sapien
258	15	53.6	420	2	Y461_STNY3	Q55137 synchocyst	331	15	53.6	716	2	Q5RDI3_PONPY	Q5rld3 pongo pygma
259	15	53.6	422	2	Q7R075_GIALA	Q7r075 giardia lam	332	15	53.6	727	2	Q9UB84_9CILI	Q9ub84 paramesio
260	15	53.6	423	2	Q9C9J4_ARATH	Q9c9j4 arabidopsis	333	15	53.6	727	2	Q8VL05_ALRNO	Q8v105 altermonas
261	15	53.6	428	2	Q4SEM6_TERNG	Q4sem6 tetraodon n	334	15	53.6	731	2	Q8RTZ9_9GAMM	Q8rtz9 pseudosalter
262	15	53.6	432	2	Q9VMD4_DROME	Q9vmd4 drosophila	335	15	53.6	743	2	Q4T862_TERNG	Q4t862 tetraodon n
263	15	53.6	433	2	Q5B5O6_EMENT	Q5b5o6 aspergillus	336	15	53.6	746	2	Q4SPR7_TERNG	Q4spr7 tetraodon n
264	15	53.6	443	2	Q8LPZ9_ARATH	Q8lpz9 arabidopsis	337	15	53.6	754	2	Q4SSB1_TERNG	Q4ssb1 tetraodon n
265	15	53.6	443	2	Q93Z67_ARATH	Q93z67 arabidopsis	338	15	53.6	771	2	Q8K4I3_MOUSE	Q8k4i3 mus musculu
266	15	53.6	459	2	Q9XXP6_CABEL	Q9xpx6 caenorhabdi	339	15	53.6	772	2	Q4WK08_ASPFU	Q4wk08 aspergillus
267	15	53.6	459	2	Q4E207_LEITWA	Q4e207 leishmania	340	15	53.6	799	2	Q5ONF5_ORYSA	Q5onf5 oryza sativ
268	15	53.6	460	2	Q725N8_DESYH	Q725n8 desulfovibr	341	15	53.6	811	2	Q91582_HPBVO	Q91582 hepatitis b
269	15	53.6	461	2	Q58317_9TRYP	Q58317 trypanosoma	342	15	53.6	813	2	Q9BLJ2_CIOIN	Q9blj2 clona intes
270	15	53.6	464	2	Q7R682_GIALA	Q7r682 giardia lam	343	15	53.6	833	2	Q6J288_ACACA	Q6j288 acanthamoeb
271	15	53.6	465	2	Q4RET0_TERNG	Q4ret0 tetraodon n	344	15	53.6	840	1	RNFI9_MOUSE	Q91563 hepatitis b
272	15	53.6	470	2	Q8X0Z2_GIBRU	Q8x0z2 gibberella	345	15	53.6	843	2	Q91572_HPBVO	Q91572 hepatitis b
273	15	53.6	471	2	Q9VMG7_DROME	Q9vmg7 drosophila	346	15	53.6	843	2	Q81141_HPBVO	Q81141 hepatitis b
274	15	53.6	485	2	Q5U795_AZOLI	Q5u795 azospirilli	347	15	53.6	843	2	Q8UXG0_HPBVO	Q8uxg0 hepatitis b
275	15	53.6	489	1	ZIF1_CABEL	P34482 caenorhabdi	348	15	53.6	869	2	Q8FGB3_ECOL6	Q8fgb3 escherichia
276	15	53.6	497	2	Q7UPM3_RHOBA	Q7upm3 rhodospirell	349	15	53.6	869	2	Q8PMF6_XANAC	Q8pmf6 xanthomonas
277	15	53.6	500	2	Q7UHJ6_RHOBA	Q7uhj6 rhodospirell	350	15	53.6	885	2	Q6R389_9PROT	Q6r389 unclutred
278	15	53.6	503	2	Q4YZM1_PLABE	Q4yzm1 plasmodium	351	15	53.6	888	2	Q8RPZ1_9PROT	Q8rpz1 drosophila
279	15	53.6	505	2	Q4FKT0_9TRYP	Q4fkt0 trypanosoma	352	15	53.6	888	2	Q81PZ1_LEITWA	Q81p21 leishmania
280	15	53.6	516	2	Q54C18_DICDI	Q54c18 dictyostell	353	15	53.6	904	2	Q4QAU0_LEITWA	Q4qau0 leishmania
281	15	53.6	516	2	Q87P76_VIBPA	Q87p76 vibrio para	354	15	53.6	935	2	Q94719_PARTE	Q94719 paramesliu
282	15	53.6	520	2	Q8N1N4_HUMAN	Q8n1n4 homo sapien	355	15	53.6	937	2	Q5N1I1_FRAFT	Q5n1i1 francisella
283	15	53.6	521	2	Q7RTT2_HUMAN	Q7rtc2 homo sapien	356	15	53.6	943	2	Q6ODG4_ORYSA	Q6odg4 oryza sativ
284	15	53.6	522	2	Q5BAW2_EMENT	Q5baw2 aspergillus	357	15	53.6	948	2	Q5NXL5_AZOSE	Q5nxl5 azocarcus sp
285	15	53.6	525	2	Q5ZQI1_CRYNE	Q5zqi1 cryptococcu	358	15	53.6	969	2	Q8BR86_MOUSE	Q8br86 mus musculu
286	15	53.6	525	2	Q5KP13_CRYNE	Q5kpl3 cryptococcu	359	15	53.6	984	2	Q81PZ2_DROME	Q81p22 drosophila
287	15	53.6	527	2	Q6ERD6_ORYSA	Q6erd6 oryza sativ	360	15	53.6	988	2	Q7QYU1_GIALA	Q7qyu1 giardia lam
288	15	53.6	531	1	PGLR1_RALSO	P58158 ralatonia s	361	15	53.6	993	2	Q9VOL8_DROME	Q9vol8 drosophila
289	15	53.6	538	2	Q51RL7_MAGNOR	Q51rl7 magnaporthe	362	15	53.6	1002	1	SPHR_AMEPV	Q54M27 dictostell
290	15	53.6	541	2	Q4WJAO_ASPFU	Q4wjao aspergillus	363	15	53.6	1003	2	Q54M27_DICDI	Q54m27 dictostell
291	15	53.6	544	2	Q941V2_ORYSA	Q941v2 oryza sativ	364	15	53.6	1013	2	Q4SR96_TERNG	Q4sr96 tetraodon n
292	15	53.6	549	2	P77818_AGRVI	Q941v2 agrobacteri	365	15	53.6	1023	2	Q5JUB6_HUMAN	Q5jue6 homo sapien
293	15	53.6	563	2	Q6V9W0_RAT	Q6v9w0 rattus norv	366	15	53.6	1025	2	Q7R6J7_GIALA	Q7r6j7 giardia lam
294	15	53.6	564	2	Q9YQZ3_9HERP	Q9yqz3 randt herpe	367	15	53.6	1025	2	Q94EN9_CHLRE	Q94en9 chlamydomon
295	15	53.6	565	1	STK23_MOUSE	Q9zqg2 mus musculu	368	15	53.6	1039	2	Q610M2_CABER	Q610m2 caenorhabdi
296	15	53.6	569	2	Q9CBD6_MYCLE	Q9cbd6 mycobacteri	369	15	53.6	1043	2	Q17644_CABEL	Q17644 caenorhabdi
297	15	53.6	574	2	Q8MXW1_HALRO	Q8mxw1 halocynthia	370	15	53.6	1057	2	Q4N4P8_THERA	Q4n4p8 thelletia p
298	15	53.6	587	2	Q9VHU6_DROME	Q9vhu6 drosophila	371	15	53.6	1061	2	Q4XQ34_PLACH	Q4xq34 plasmodium
299	15	53.6	589	2	Q4P936_USTWA	Q4p936 ustilago ma	372	15	53.6	1063	2	Q7QU10_GIALA	Q7qu10 giardia lam
300	15	53.6	593	2	Q818V8_GIALA	Q818v8 giardia lam	373	15	53.6	1068	2	Q4QEB6_LEITWA	Q4qeb6 leishmania
301	15	53.6	593	2	Q8LMC6_ORYSA	Q8lmc6 oryza sativ	374	15	53.6	1070	2	Q7R2W4_GIALA	Q7r2w4 giardia lam
302	15	53.6	595	2	Q5VXL7_HUMAN	Q5vxl7 homo sapien	375	15	53.6	1089	2	Q4Q9C2_LEITWA	Q4q9c2 leishmania
303	15	53.6	599	2	Q8W4D7_ARATH	Q8w4d7 arabidopsis	376	15	53.6	1103	2	Q8WR15_DROME	Q8wr15 drosophila
304	15	53.6	602	2	Q6MG12_NEUCR	Q6mg12 neucspora	377	15	53.6	1146	2	Q7OX15_GIALA	Q7ox15 giardia lam
305	15	53.6	605	2	Q9SFU8_ARATH	Q9sfu8 arabidopsis	378	15	53.6	1173	2	Q5U4V9_XENIA	Q5u4v9 xenopus lae
306	15	53.6	608	2	Q7SDR2_NEUCR	Q7sdr2 neucspora	379	15	53.6	1186	1	C14AA_BACTE	Q45710 bacillus th
307	15	53.6	609	2	Q64BX6_9ARCH	Q64bx6 unclutred	380	15	53.6	1189	2	Q4NUU7_9DELT	Q4nuu7 anaeromoxb
308	15	53.6	614	2	Q64DG6_9ARCH	Q64dg6 unclutred	381	15	53.6	1191	2	Q8RL60_PSEPT	Q8rl60 pseudomonas
309	15	53.6	614	2	Q8GZ21_ARATH	Q8gz21 arabidopsis	382	15	53.6	1216	2	Q4PDB3_USTWA	Q4pdb3 ustilago ma
310	15	53.6	622	2	Q8C9V4_MOUSE	Q8c9v4 mus musculu	383	15	53.6	1232	2	Q4T1X3_TERNG	Q4t1x3 tetraodon n
311	15	53.6	633	2	Q818W0_GIALA	Q818w0 giardia lam	384	15	53.6	1236	2	Q8OTKO_MOUSE	Q8otko mus musculu
312	15	53.6	635	2	Q7QBP07_GIALA	Q7qbp07 giardia lam	385	15	53.6	1249	2	Q57VB8_9TRYP	Q57vb8 trypanosoma
313	15	53.6	637	2	Q4UIV3_THERA	Q4uiv3 thelletia a	386	15	53.6	1253	2	Q4T0S1_TERNG	Q4t0s1 tetraodon n
314	15	53.6	638	2	Q7QOC4_GIALA	Q7qoc4 giardia lam	387	15	53.6	1285	2	Q8DA47_VIBVU	Q8da47 vibrio vuln
315	15	53.6	639	2	Q8PMW5_METWA	Q8pmw5 methanosarc	388	15	53.6	1297	2	Q7MK32_VIBVY	Q7mk32 vibrio vuln
316	15	53.6	640	2	Q8ITZ6_METAC	Q8itz6 methanosarc	389	15	53.6	1324	2	Q8SYK2_DROME	Q8syk2 drosophila
317	15	53.6	640	2	Q8CI21_MOUSE	Q8ci21 mus musculu	390	15	53.6	1324	2	Q9VYTR8_DROME	Q9vtr8 drosophila
318	15	53.6	644	2	Q7QWT5_GIALA	Q7qwt5 giardia lam	391	15	53.6	1343	2	Q4IK63_GIBZE	Q4ik63 gibberella
319	15	53.6	659	2	Q6C9W2_YARLI	Q6c9w2 yarrowia.l1	392	15	53.6	1353	2	Q6R2R5_ARATH	Q6r2r5 arabidopsis
320	15	53.6	659	2	Q4NGR1_THERA	Q4ngr1 thelletia p	393	15	53.6	1361	2	Q9NGCV2_DROME	Q9ngcv2 drosophila
321	15	53.6	669	2	Q4MMZ9_ASPFU	Q4mmz9 aspergillus	394	15	53.6	1361	2	Q9V7T14_DROME	Q9v7t14 drosophila
322	15	53.6	679	2	Q6PUC7_ANOCA	Q6puc7 anopheles g	395	15	53.6	1364	2	Q9SW45_ARATH	Q9sw45 arabidopsis
323	15	53.6	688	2	Q96VF2_USTWA	Q96vf2 ustilago ma	396	15	53.6	1376	2	Q8X1P2_PODAN	Q8x1p2 podatospora a

397	15	53.6	1414	2	040E45_LEITMA	04q645 leishmania	470	15	53.6	10495	2	04RE92_TETNG	04re92 tetraodon n
398	15	53.6	1429	2	07P0W3_ANOGA	07p0w3 anopheles g	471	15	53.6	34350	2	08WZ42_HUMAN	08wz42 homo sapien
399	15	53.6	1447	2	04WPJ3_ASPEPU	04wpj3 aspergillus	472	15	50.0		2	07WZ76_GASCO	07wz76 gaestricoloni
400	15	53.6	1451	1	EM30_ARATH	Q42510 arabidopsis	473	14	50.0	21	2	05ZD65_MAGAR	05zd65 magnaporthe
401	15	53.6	1468	2	080TF6_MOUSE	080tf6 mus musculu	474	14	50.0	29	2	04TIR3_ANGA	04tir3 tetraodon n
402	15	53.6	1514	2	04FX85_LEIMA	04fx85 leishmania	475	14	50.0	49	2	05YTU2_ANOGA	05ytu2 anopheles g
403	15	53.6	1516	2	07RLJ5_PLAYO	07rlj5 plasmodium	476	14	50.0	46	2	07RXU9_NEUCR	07rxu9 neucspora
404	15	53.6	1533	2	07RLI0_PLAYO	07rli0 plasmodium	477	14	50.0	52	2	07JFK9_RHOBA	07jfk9 rhodospirell
405	15	53.6	1533	2	04Q5Q2_LEIMA	04q5q2 leishmania	478	14	50.0	53	2	07NS59_RAT	07ns59 rattus norv
406	15	53.6	1539	2	07RT03_PLAYO	07rt03 plasmodium	479	14	50.0	55	2	05VIZ4_AERHY	05viz4 aetromonas h
407	15	53.6	1624	2	07RM69_PLAYO	07rm69 plasmodium	480	14	50.0	57	2	04TIC2_ANGA	04tic2 tetraodon n
408	15	53.6	1625	2	04T364_TETNG	04t364 tetraodon n	481	14	50.0	58	2	05EY69_ADE07	05ey69 human adeno
409	15	53.6	1651	2	07RNX5_PLAYO	07rnx5 plasmodium	482	14	50.0	64	2	04YVW8_PLABE	04yvwm plasmodium
410	15	53.6	1722	2	04NOK1_GDELT	04nkl1 anaeromyxob	483	14	50.0	70	2	04YBP5_PLABE	04ybps plasmodium
411	15	53.6	1729	2	07XGP6_ORYSA	07xgp6 oryza sativ	484	14	50.0	72	2	084Z47_ORYSA	084z47 oryza sativ
412	15	53.6	1729	2	08RLZ0_ORYSA	08rlz0 oryza sativ	485	14	50.0	73	2	08RV11_PINPS	08rv11 pinus pins
413	15	53.6	1755	2	06RKI6_BOTCI	06rk16 botrytis ci	486	14	50.0	74	2	087AK6_XYLFT	087ak6 xyella fas
414	15	53.6	1782	2	04MQZ5_ASPEPU	04mqz5 aspergillus	487	14	50.0	76	2	05TYD0_ANOGA	05tyd0 anopheles g
415	15	53.6	1782	2	06X012_SOLIN	06x012 solenopsis	488	14	50.0	77	2	089W35_BRAJA	089w35 bradyrhizob
416	15	53.6	1783	2	04FMM9_LEIMA	04fmm9 leishmania	489	14	50.0	77	2	09T0Z3_9CAUD	09t0z3 bacteriopho
417	15	53.6	1806	2	05BH30_EMENT	05bh30 aspergillus	490	14	50.0	78	2	07UHI8_RHOBA	07uhi8 rhodospirell
418	15	53.6	1893	2	04U0S1_BRABE	04u0s1 brachydanio	491	14	50.0	80	2	06TNI7_STYTH	06tni7 symbiodace
419	15	53.6	1944	2	06MS81_9ACTO	06ms81 actinomadura	492	14	50.0	80	2	04T6X9_TETNG	04t6x9 tetraodon n
420	15	53.6	1946	2	054YA3_DICDI	054ia3 dictyostell	493	14	50.0	82	2	04S1R2_TETNG	04s1r2 tetraodon n
421	15	53.6	1958	2	084HN4_9ACTO	084hn4 kitasatospo	494	14	50.0	83	2	069MR0_ORYSA	069mr0 oryza sativ
422	15	53.6	1961	2	084HM9_9ACTO	084hm9 streptomyce	495	14	50.0	86	2	05P1S6_AZOSE	05p1s6 azaracus sp
423	15	53.6	1976	2	07PUM1_ANOGA	07pum1 anopheles g	496	14	50.0	88	2	0970Z3_BREIN	0970z3 brevibacter
424	15	53.6	1977	2	08WU72_ANOGA	08wu72 anopheles g	497	14	50.0	88	2	09PFI7_XYLFA	09pfi7 xyella fas
425	15	53.6	1978	2	08T5H2_ANOGA	08t5h2 anopheles g	498	14	50.0	90	2	057D56_BRUAB	057d56 bruceella ab
426	15	53.6	1988	2	06RKI8_BOTCI	06rk18 botrytis ci	499	14	50.0	91	2	04XC49_PLACH	04xc49 plasmidum
427	15	53.6	1996	2	04T5Z8_TETNG	04t5z8 tetraodon n	500	14	50.0	91	2	091667_XENIA	091667 xenopus lae
428	15	53.6	2043	2	04Q510_LEIMA	04q510 leishmania	501	14	50.0	93	1	RS16_CAMP	08kx84 candidatus
429	15	53.6	2067	2	04S488_TETNG	04s488 tetraodon n	502	14	50.0	93	2	087Z87_NEUCR	087z87 neucspora
430	15	53.6	2088	2	P91443_CAEBL	P91443 caenorhabdi	503	14	50.0	94	2	07PC59_MYCGM	07pc59 mycobacteri
431	15	53.6	2099	2	06Z2K8_CAEBR	06z2k8 caenorhabdi	504	14	50.0	94	2	P91216_CAEBL	P91216 caenorhabdi
432	15	53.6	2106	2	06RWD9_NEUCR	06rwd9 neucspora	505	14	50.0	97	2	053CUI_9GAMA	053cul macaca fusc
433	15	53.6	2112	2	08WPL0_9UROC	08wpl0 oikopleura	506	14	50.0	97	2	07S7T5_NEUCR	07s7t5 neucspora
434	15	53.6	2117	2	074H49_9GOST	074h49 geobacter s	507	14	50.0	98	2	07YFR0_NEUCR	07yfr0 neucspora
435	15	53.6	2159	2	09Y8A6_9PEZI	09y8a6 nodulipor	508	14	50.0	100	2	09DXR8_9CTNC	09dxr8 beak and fe
436	15	53.6	2159	2	086T05_DICDI	086t05 dictyostell	509	14	50.0	101	2	09DXG4_9CTNC	09dxg4 beak and fe
437	15	53.6	2159	2	0551N7_DICDI	0551n7 dictyostell	510	14	50.0	101	2	09DXG7_9CTNC	09dxg7 beak and fe
438	15	53.6	2162	2	08NR46_9PEZI	08nr46 xylaria sp.	511	14	50.0	102	2	09WRM6_GRALE	09wrm6 giclarilla
439	15	53.6	2162	2	09VOM0_DROME	09vom0 dirosophila	512	14	50.0	102	2	09WRM6_GRALE	09wrm6 giclarilla
440	15	53.6	2173	2	051PP8_MAGGR	051pp8 magnaporthe	513	14	50.0	102	2	08Z4W5_CHLGV	08z4w5 chlamydom
441	15	53.6	2176	2	046112_DROME	046112 dirosophila	514	14	50.0	103	1	Y040_MYCPA	P75074 mycoplasma
442	15	53.6	2181	1	STCA_EMENT	Q12397 emeritella	515	14	50.0	103	2	05YF63_9YTRU	05yf63 rock bream
443	15	53.6	2187	2	P79068_GLOIA	P79068 glomerella	516	14	50.0	104	2	013588_YEAST	013588 saccharomyc
444	15	53.6	2187	2	08TGD7_ASPEPE	08tgd7 aspergillus	517	14	50.0	105	2	04TPW2_9EPHN	04tpw2 erythrobact
445	15	53.6	2188	2	06XR12_9PEZI	06xr12 aspergillus	518	14	50.0	107	1	YB56_YEAST	04c6r2 saccharomyc
446	15	53.6	2203	2	07S7J6_NEUCR	07s7j6 neucspora	519	14	50.0	107	2	007521_YEAST	007521 saccharomyc
447	15	53.6	2238	2	05AIV5_EMENT	05aiv5 aspergillus	520	14	50.0	108	2	06K452_ORYSA	06k452 oryza sativ
448	15	53.6	2275	2	06RKE4_COCHR	06rke4 cochlloboln	521	14	50.0	109	2	082YH0_STRAW	082yh0 streptomyc
449	15	53.6	2397	2	06Q7Y4_PAPRP	06q7y4 parametium	522	14	50.0	110	2	07SC86_NEUCR	07sc86 neucspora
450	15	53.6	2511	2	04NKK5_9DELT	04nkk5 anaeromyxob	523	14	50.0	111	2	08IX35_HUMAN	08ix35 homo sapien
451	15	53.6	2517	2	05AKZ3_EMENT	05akz3 aspergillus	524	14	50.0	111	2	P72474_STRMU	P72474 streptococc
452	15	53.6	2533	2	P90589_TAPPE	P90589 parametium	525	14	50.0	112	2	051906_DESVU	051906 deusilfovbr
453	15	53.6	2533	2	027183_TAPPE	027183 parametium	526	14	50.0	112	2	04V572_DROME	04v572 dirosophila
454	15	53.6	2543	2	P90649_PAPRP	P90649 parametium	527	14	50.0	113	2	04V572_DROME	04v572 dirosophila
455	15	53.6	2556	2	07QJEB3_ANOGA	07qjeb3 anopheles g	528	14	50.0	113	2	09ZSE1_PINKA	09zse1 pinus radia
456	15	53.6	2644	2	06RKI2_BOTCI	06rk12 botrytis ci	529	14	50.0	115	2	056754_9HBPC	056754 heparitis c
457	15	53.6	2703	1	NOTCH_DROME	P07207 dirosophila	530	14	50.0	115	2	0976E2_SULTO	0976e2 sulfolobus
458	15	53.6	2717	2	094710_PAPPE	094710 parametium	531	14	50.0	116	2	08W633_9CAUD	08w633 bacteriopho
459	15	53.6	2729	2	06P0K6_PAPPE	06p0k6 parametium	532	14	50.0	118	2	0868W8_PLAYA	0868w8 plasmodium
460	15	53.6	2773	2	0591S5_BRABE	0591s5 brachydanio	533	14	50.0	118	2	08RPE3_FUSNN	08rpe3 fusobacteri
461	15	53.6	2793	2	05BBJ6_EMENT	05bbj6 aspergillus	534	14	50.0	119	2	04QAH0_LEIMA	04qah0 leishmania
462	15	53.6	2836	2	09V8L8_DROME	09v8l8 dirosophila	535	14	50.0	120	2	051901_DESBA	051901 deusilfovbr
463	15	53.6	4039	2	04KCD6_PSEPS	04kcd6 pseudomonas	536	14	50.0	120	2	07P6A3_FUSNV	07p6a3 fusobacteri
464	15	53.6	4376	2	05MP07_9BACT	05mp07 symbiotin ba	537	14	50.0	121	2	084ZU5_ORYSA	084zu5 oryza sativ
465	15	53.6	4767	2	017301_CAEBR	017301 caenorhabdi	538	14	50.0	121	2	04SZW6_TETNG	04szw6 tetraodon n
466	15	53.6	4881	2	06TAB6_9BACT	06tab6 symbiotin ba	539	14	50.0	122	2	069HS6_CTOIN	069hs6 ciona intes
467	15	53.6	7743	2	0618F4_CAEBR	0618f4 caenorhabdi	540	14	50.0	123	2	06T9W1_NICSY	06t9w1 nicotiana s
468	15	53.6	7829	2	018559_CAEBL	018559 caenorhabdi	541	14	50.0	123	2	051904_9DELT	051904 deusilfovbr
469	15	53.6	7962	2	Q10465_HUMAN	Q10465 homo sapien	542	14	50.0	123	2	08VXR8_MYCTU	08vxr8 mycobacteri

543	14	50.0	123	2	093464_CABAU	093464	carassius a	616	14	50.0	158	2	0641H1_9CIRC	0641H1	beak and fe
544	14	50.0	124	2	07S9A8_NEUCR	07S9A8	neupora	617	14	50.0	158	2	09DXF2_9CIRC	09DXF2	beak and fe
545	14	50.0	124	2	08NB52_HUMAN	08NB52	homo sapien	618	14	50.0	158	2	09DXF5_9CIRC	09DXF5	beak and fe
546	14	50.0	124	2	051916_9DELT	051916	unidentifie	619	14	50.0	159	2	06RZ20_9PEPT	06RZ20	periconia m
547	14	50.0	124	2	04V2P3_BURMA	04V2P3	burholderi	620	14	50.0	159	2	09PDM1_9PEPT	09PDM1	xyliella fas
548	14	50.0	125	2	06E1W5_9EURO	06E1W5	petromyces	621	14	50.0	160	2	04SYU6_TETNG	04SYU6	tetradon n
549	14	50.0	125	2	06E1W8_9EURO	06E1W8	aspergillus	622	14	50.0	161	2	06ZUT2_HUMAN	06ZUT2	homo sapien
550	14	50.0	125	2	06E1W9_9EURO	06E1W9	petromyces	623	14	50.0	161	2	09S178_MACMU	09S178	macaca mula
551	14	50.0	125	2	06E1X0_9EURO	06E1X0	petromyces	624	14	50.0	162	2	04IGT6_ADEP3	04IGT6	porcine ade
552	14	50.0	125	2	07R1V6_GIALA	07R1V6	giardia lam	625	14	50.0	162	2	04TH15_TETNG	04TH15	tetradon n
553	14	50.0	126	2	09JH18_MOUSE	09JH18	mus musculus	626	14	50.0	163	2	06RZ18_9EURO	06RZ18	talarmyces
554	14	50.0	129	1	LYSCI1_CANFA	LYSCI1	canis famli	627	14	50.0	163	2	09P280_HUMAN	09P280	homo sapien
555	14	50.0	129	1	05C4S6_SCHJA	05C4S6	schistosoma	628	14	50.0	163	2	0967R9_GIALA	0967R9	giardia lam
556	14	50.0	130	2	06RZ24_9PLEO	06RZ24	stemphylium	629	14	50.0	163	2	069X90_ORYSA	069X90	oryza sativ
557	14	50.0	131	1	NEUZ ANSAN	PI9630	anser anser	630	14	50.0	164	2	06TGO1_9LECA	06TGO1	lecanora ma
558	14	50.0	132	2	07S1O7_NEUCR	07S1O7	neupospora	631	14	50.0	164	2	06TGO0_9LECA	06TGO0	lecanora na
559	14	50.0	132	2	09VLJ37_DROME	09VLJ37	drosophila	632	14	50.0	164	2	06TGO2_9LECA	06TGO2	lecanora fl
560	14	50.0	132	2	08G4M6_BIFLO	08G4M6	bifidobacte	633	14	50.0	164	2	06TGO3_9PEPT	06TGO3	dermatocarp
561	14	50.0	133	1	YB9V YEAST	P38350	saccharomyc	634	14	50.0	164	2	044586_CABEL	044586	caenorhabd
562	14	50.0	133	2	0870B4_9PUNG	0870B4	picomyces s	635	14	50.0	164	2	05SNE2_ORYSA	05SNE2	oryza sativ
563	14	50.0	133	2	05TVJ5_ANOGA	05TVJ5	anopheles g	636	14	50.0	165	2	044587_CABEL	044587	caenorhabd
564	14	50.0	134	2	08GTU1_HUMAN	08GTU1	homo sapien	637	14	50.0	166	2	09BHK2_9ANNE	09BHK2	sabella spa
565	14	50.0	134	2	06ZFU3_ORYSA	06ZFU3	oryza sativ	638	14	50.0	167	2	05H5G2_XANOR	05H5G2	xanthomonas
566	14	50.0	135	2	05Q994_IXOSC	05Q994	ixodes scap	639	14	50.0	167	2	08YK95_MYCTU	08YK95	mycobacteri
567	14	50.0	135	2	018158_CABEL	018158	caenorhabd	640	14	50.0	168	2	08H2J1_ORYSA	08H2J1	oryza sativ
568	14	50.0	135	2	092P10_RHIME	092P10	rhizobium m	641	14	50.0	169	2	06REU4_9FLOR	06REU4	halymenia s
569	14	50.0	136	2	06TIL3_DROME	06TIL3	drosophila	642	14	50.0	169	2	082G98_STRAW	082G98	streptomyces
570	14	50.0	137	2	08TUJ9_METAC	08TUJ9	methanosarc	643	14	50.0	169	2	06ZMG5_HUMAN	06ZMG5	homo sapien
571	14	50.0	137	2	07Q4X7_ANOGA	07Q4X7	anopheles g	644	14	50.0	170	2	07UZZ4_RHOBA	07UZZ4	rhodopirell
572	14	50.0	137	2	08BWD7_MOUSE	08BWD7	m mus muscu	645	14	50.0	170	2	08LNA4_ORYSA	08LNA4	oryza sativ
573	14	50.0	139	2	06RZ21_9EURO	06RZ21	dichotomomy	646	14	50.0	171	2	091B04_9ACTO	091B04	oryza sativ
574	14	50.0	139	2	000830_LEIMA	000830	leishmania	647	14	50.0	171	2	08S2V2_DROME	08S2V2	drosophila
575	14	50.0	141	2	077076_HARCO	077076	haemophilus	648	14	50.0	172	2	05B025_EMENTI	05B025	aspergillus
576	14	50.0	141	2	084MF8_ORYSA	084MF8	oryza sativ	649	14	50.0	174	2	06RZ25_9PLEO	06RZ25	westernykel
577	14	50.0	143	1	PONA_DICTDI	P54660	dicyoscell	650	14	50.0	174	2	081PD4_DROME	081PD4	drosophila
578	14	50.0	143	2	0585A5_9ITRP	0585A5	trypanosoma	651	14	50.0	175	1	SNAC_STRPR	SNAC	streptomyces
579	14	50.0	143	2	054BG4_DICTDI	054BG4	dicyoscell	652	14	50.0	175	2	05STU7_CRYNE	05STU7	cryptococcu
580	14	50.0	143	2	06NA71_RHOBA	06NA71	rhodopseudo	653	14	50.0	175	2	08GN71_BIFLO	08GN71	bifidobacte
581	14	50.0	145	2	07R3A4_GIALA	07R3A4	giardia lam	654	14	50.0	175	2	05H2G2_XANOR	05H2G2	xanthomonas
582	14	50.0	148	2	052CL3_MAGGR	052CL3	magnaporthe	655	14	50.0	175	2	07WX94_ALCEU	07WX94	alecaligenes
583	14	50.0	148	2	08E2V5_MYCPE	08E2V5	mycoplasma	656	14	50.0	175	2	082PBE_STRAW	082PBE	streptomyces
584	14	50.0	149	2	06RZ22_9PEPT	06RZ22	cladosporiu	657	14	50.0	175	2	09R2P6_STRAW	09R2P6	pseudonotri
585	14	50.0	149	2	07QY01_ANOGA	07QY01	anopheles g	658	14	50.0	176	2	06RZ19_9ASCO	06RZ19	polyisiphoni
586	14	50.0	149	2	07XG18_ORYSA	07XG18	oryza sativ	659	14	50.0	177	1	PHEB_POLBO	PHEB	porphyra te
587	14	50.0	149	2	094T02_ORYSA	094T02	oryza sativ	660	14	50.0	177	1	PHEB_PORPU	PHEB	porphyra te
588	14	50.0	149	2	08AV79_BRARE	08AV79	brachydanio	661	14	50.0	177	1	PHEB_PORPE	PHEB	porphyra ye
589	14	50.0	150	2	06RZ27_9EURO	06RZ27	penicillium	662	14	50.0	177	1	09YD43_AERPE	09YD43	aeropyrum p
590	14	50.0	150	2	06TGP9_9LECA	06TGP9	lecanora et	663	14	50.0	177	2	04WUY6_ASPFU	04WUY6	aspergillus
591	14	50.0	150	2	061641_ONCOC	061641	onchocerca	664	14	50.0	177	2	08NAA9_HUMAN	08NAA9	homo sapien
592	14	50.0	150	2	05NAT9_ORYSA	05NAT9	oryza sativ	665	14	50.0	177	2	09MRW8_GRALE	09MRW8	gracilaria
593	14	50.0	151	2	07PKW7_ANOGA	07PKW7	anopheles g	666	14	50.0	177	2	07S1P0_GRIAR	07S1P0	gracilaria
594	14	50.0	152	2	06RZ15_ASPNG	06RZ15	aspergillus	667	14	50.0	177	2	068FM1_MOUSE	068FM1	mus musculu
595	14	50.0	154	2	09YAS6_AERPE	09YAS6	aeropyrum p	668	14	50.0	178	2	09UAV9_CABEL	09UAV9	caenorhabd
596	14	50.0	155	2	0857R5_9CAUD	0857R5	mycobacteri	669	14	50.0	178	2	060MW3_CABER	060MW3	caenorhabd
597	14	50.0	156	1	PPDA_ECOLI	P33554	escherichia	670	14	50.0	179	2	067328_9INFA	067328	influenza a
598	14	50.0	156	2	016422_CABEL	016422	caenorhabd	671	14	50.0	179	2	07YT17_CABEL	07YT17	caenorhabd
599	14	50.0	156	2	06SEV4_9BACT	06SEV4	uncultured	672	14	50.0	181	2	04LZX3_9BTRK	04LZX3	burholderi
600	14	50.0	156	2	057KB8_SALCH	057KB8	salmonella	673	14	50.0	181	2	08C8W0_MOUSE	08C8W0	mus musculu
601	14	50.0	156	2	05PEM2_SALPA	05PEM2	salmonella	674	14	50.0	182	2	04T1Y1_CABEL	04T1Y1	caenorhabd
602	14	50.0	156	2	08ZMB0_SALTY	08ZMB0	salmonella	675	14	50.0	183	2	069T53_ORYSA	069T53	oryza sativ
603	14	50.0	156	2	082413_SALTY	082413	salmonella	676	14	50.0	184	2	08H447_ORYSA	08H447	oryza sativ
604	14	50.0	157	2	025068_HAEIR	025068	haemotobia	677	14	50.0	184	2	05TMH6_ANOGA	05TMH6	anopheles g
605	14	50.0	157	2	06XHV2_DROYA	06XHV2	drosophila	678	14	50.0	186	2	DUS18_BOVIN	DUS18	bos taurus
606	14	50.0	157	2	09BH65_GIALA	09BH65	giardia lam	679	14	50.0	188	1	DUS18_HUMAN	DUS18	homo sapien
607	14	50.0	157	2	09B1L6_GIALA	09B1L6	giardia lam	680	14	50.0	188	1	DUS18_MOUSE	DUS18	mus musculu
608	14	50.0	157	2	06AD65_LEIXX	06AD65	leishmania x	681	14	50.0	188	1	DUS18_PONPY	DUS18	pongo pygma
609	14	50.0	158	2	006775_HAEIR	006775	haemotobia	682	14	50.0	188	2	04R3T3_MACFA	04R3T3	macaca fasc
610	14	50.0	158	2	09B1L8_GIALA	09B1L8	giardia lam	683	14	50.0	189	2	0581P6_9CAUD	0581P6	cyanoceph
611	14	50.0	158	2	06S8B2_BACID	06S8B2	bacillus i	684	14	50.0	189	2	09D9D8_MOUSE	09D9D8	mus musculu
612	14	50.0	158	2	090236_9CIRC	090236	beak and fe	685	14	50.0	189	2	052046_9ZZZZ	052046	plasmid phy
613	14	50.0	158	2	0641F6_9CIRC	0641F6	beak and fe	686	14	50.0	189	2	087TJ5_VIBPA	087TJ5	vibrio para
614	14	50.0	158	2	0641G0_9CIRC	0641G0	beak and fe	687	14	50.0	190	2	04MSA1_BACCE	04MSA1	bacillus ce
615	14	50.0	158	2	0641G7_9CIRC	0641G7	beak and fe	688	14	50.0	191	2			



689	14	50.0	193	2	Q4PD06_USMTMA	Q4PD06_usfllago_ma	762	14	50.0	221	1	QAZ1_MESAU	P70112_mesocricetu
690	14	50.0	193	2	Q5TV08_ANOGA	Q5TV08_anopheles_g	763	14	50.0	221	2	Q4QSJ7_9LECA	Q4qsj7_pertuaria
691	14	50.0	193	2	Q9EM12_9CCRO	Q9em12_avian_infec	764	14	50.0	221	2	Q8XEN5_CHUTE	Q8xens_chlorobium
692	14	50.0	195	2	Q54199_STREGR	Q54199_streptomyce	765	14	50.0	221	2	Q5XJN5_BRABE	Q5xjns_brachydano
693	14	50.0	197	2	Q4WFH8_ASPFU	Q4wfh8_aspergillus	766	14	50.0	222	2	Q94722_PARTE	Q94722_paramecium
694	14	50.0	197	2	Q7R0J0_GIALFA	Q7r0j0_giardia_lam	767	14	50.0	222	2	Q56PH2_RUEDE	Q56ph2_ruegeria_ge
695	14	50.0	197	2	Q44585_CAEEL	Q44585_caenorhabdi	768	14	50.0	224	2	Q56P81_9METZ	Q56p81_uncultured
696	14	50.0	197	2	Q8GGS6_STRAZ	Q8gg86_streptomyce	769	14	50.0	225	2	Q20590_CABETL	Q20590_caenorhabdi
697	14	50.0	198	2	DUS14_HUMAN	Q95147_homo_sapien	770	14	50.0	226	2	Q5PC56_SALPA	Q5pc56_salmonella
698	14	50.0	198	1	DUS14_MOUSE	Q911y7_homo_sapien	771	14	50.0	226	2	Q4RDM1_TETNG	Q4rdm1_tetradodon_n
699	14	50.0	198	2	Q6FI36_HUMAN	Q6fi36_homo_sapien	772	14	50.0	227	2	Q4QSG1_9LECA	Q4qsg1_pertuaria
700	14	50.0	198	2	Q8VKK4_MYCTU	Q8vkk4_mycobacteri	773	14	50.0	227	2	Q4QSI1_9LECA	Q4qsi1_pertuaria
701	14	50.0	198	2	Q8SV91_MOUSE	Q8sv91_mus_musculu	774	14	50.0	227	2	Q4QSI1_9LECA	Q4qsi1_pertuaria
702	14	50.0	198	2	Q7TPY1_MOUSE	Q7tpy1_mus_musculu	775	14	50.0	227	2	Q4QSN0_9ASCO	Q4qsn0_coccotrema
703	14	50.0	199	2	NIT1_EAST	P40447_saccharomyc	776	14	50.0	227	2	Q4QSN9_9ASCO	Q4qsn9_coccotrema
704	14	50.0	199	1	Q05729_YEAST	Q05129_saccharomyc	777	14	50.0	227	2	Q4QSW8_9ASCO	Q4qsw8_coccotrema
705	14	50.0	199	2	Q6Q5A1_YEAST	Q6q5a1_saccharomyc	778	14	50.0	227	2	Q4QSM7_9ASCO	Q4qsm7_coccotrema
706	14	50.0	201	2	Q7PUG4_ANOGA	Q7pug4_anopheles_g	779	14	50.0	227	2	Q4QSM4_9LECA	Q4qsm4_coccotrema
707	14	50.0	201	2	Q6DGS6_BRARE	Q6dgs6_brachydano	780	14	50.0	227	2	Q4QSM3_9LECA	Q4qsm3_coccotrema
708	14	50.0	202	2	Q4QSU5_9LECA	Q4qsj5_pertuaria	781	14	50.0	227	2	Q4QSU8_9LECA	Q4qsl8_octrolochia
709	14	50.0	202	2	Q4Q3M4_LEIMA	Q4q3m4_leishmania	782	14	50.0	227	2	Q4QSU5_9LECA	Q4qsl5_octrolochia
710	14	50.0	202	2	Q5H3M5_XANOR	Q5h3m5_xanthomonas	783	14	50.0	227	2	Q4QSL4_9LECA	Q4qsl4_pertuaria
711	14	50.0	203	2	Q4QEG4_LEIMA	Q4qeg4_leishmania	784	14	50.0	227	2	Q4QSL3_9LECA	Q4qsl3_pertuaria
712	14	50.0	203	2	Q4SFK7_TETNG	Q4sfk7_tetradodon_n	785	14	50.0	227	2	Q4QSL1_9LECA	Q4qsl1_pertuaria
713	14	50.0	204	1	DUS18_RAT	Q6axw7_rattus_novr	786	14	50.0	227	2	Q4QSK8_9LECA	Q4qsk8_pertuaria
714	14	50.0	204	2	Q5B220_EMENTI	Q5b220_aspergillus	787	14	50.0	227	2	Q4QSK6_9LECA	Q4qsk6_pertuaria
715	14	50.0	204	2	Q92472_NPYBM	Q92472_bombyx_mori	788	14	50.0	227	2	Q4QSK4_9LECA	Q4qsk4_pertuaria
716	14	50.0	205	1	Y115_NPYOP	Q10354_oryzia_pseu	789	14	50.0	227	2	Q4QSK2_9LECA	Q4qsk2_pertuaria
717	14	50.0	205	2	Q94FA0_MIMGU	Q94fa0_minulus_gut	790	14	50.0	227	2	Q4QSU6_9LECA	Q4qsj6_pertuaria
718	14	50.0	206	2	Q7Q606_ANOGA	Q7q606_anopheles_g	791	14	50.0	227	2	Q4QSU4_9LECA	Q4qsj4_pertuaria
719	14	50.0	206	2	Q8IHC4_DROME	Q8ihc4_drosophila	792	14	50.0	227	2	Q4QSU2_9LECA	Q4qsj2_pertuaria
720	14	50.0	206	2	Q81934_DROME	Q81934_drosophila	793	14	50.0	227	2	Q4QSI9_9LECA	Q4qsi9_pertuaria
721	14	50.0	206	2	Q6ZHF1_ORYSA	Q6zifi1_oryza_sativ	794	14	50.0	227	2	Q4QSI5_9LECA	Q4qsi5_pertuaria
722	14	50.0	206	2	Q4SKY8_TETNG	Q4sky8_tetradodon_n	795	14	50.0	227	2	Q4QSI4_9LECA	Q4qsi4_pertuaria
723	14	50.0	206	2	Q89183_9RETR	Q89183_feline_immu	796	14	50.0	227	2	Q4QSI3_9LECA	Q4qsi3_pertuaria
724	14	50.0	207	2	Q4QSK0_9LECA	Q4qsk0_pertuaria	797	14	50.0	227	2	Q4QSI2_9LECA	Q4qsi2_pertuaria
725	14	50.0	209	2	Q94828_TETTH	Q94828_tetrahymena	798	14	50.0	227	2	Q4QSI0_9LECA	Q4qsi0_pertuaria
726	14	50.0	209	2	Q4NDI1_9MTC	Q4ndi1_archibacte	799	14	50.0	227	2	Q4QSH9_9LECA	Q4qsh9_pertuaria
727	14	50.0	210	1	PRDX5_MOUSE	P99029_m_peroxired	800	14	50.0	227	2	Q4QSH7_9LECA	Q4qsh7_pertuaria
728	14	50.0	210	2	Q4QSH4_9LECA	Q4qsh4_pertuaria	801	14	50.0	227	2	Q4QSH3_9LECA	Q4qsh3_pertuaria
729	14	50.0	210	2	Q4X823_PLACH	Q4x823_plasmodium	802	14	50.0	227	2	Q4QSH2_9LECA	Q4qsh2_pertuaria
730	14	50.0	210	2	Q7TLR3_NPYCF	Q7tlr3_choriostonu	803	14	50.0	227	2	Q4QSH0_9LECA	Q4qsh0_pertuaria
731	14	50.0	211	2	Q4QSL7_9LECA	Q4qsl7_ochrolochia	804	14	50.0	227	2	Q4QSG7_9LECA	Q4qsg7_pertuaria
732	14	50.0	211	2	Q9D6K2_MOUSE	Q9d6k2_mus_musculu	805	14	50.0	227	2	Q4QSG5_9LECA	Q4qsg5_pertuaria
733	14	50.0	212	2	Q4QSM5_9ASCO	Q4qsm5_coccotrema	806	14	50.0	227	2	Q4QSG4_9LECA	Q4qsg4_pertuaria
734	14	50.0	212	2	Q4QSH5_9LECA	Q4qsh5_pertuaria	807	14	50.0	227	2	Q4QSG2_9LECA	Q4qsg2_pertuaria
735	14	50.0	212	2	Q5LA82_BACFN	Q5la82_bacteroides	808	14	50.0	227	2	Q4QSP8_9LECA	Q4qsp8_pertuaria
736	14	50.0	212	2	Q64OK1_BACFR	Q64ok1_bacteroides	809	14	50.0	227	2	Q4QSP4_9LECA	Q4qsp4_pertuaria
737	14	50.0	213	2	Q4QSK5_9LECA	Q4qsk5_pertuaria	810	14	50.0	227	2	Q4QSP3_9LECA	Q4qsp3_pertuaria
738	14	50.0	213	2	Q4QSH1_9LECA	Q4qsh1_pertuaria	811	14	50.0	227	2	Q4QSP1_9LECA	Q4qsp1_pertuaria
739	14	50.0	213	2	Q4QSP0_9LECA	Q4qsp0_pertuaria	812	14	50.0	227	2	Q4QSE8_9LECA	Q4qse8_varicellari
740	14	50.0	213	2	Q4P5K8_USTMA	Q4p5k8_pertuaria	813	14	50.0	227	2	Q4QSE6_9LECA	Q4qse6_varicellari
741	14	50.0	213	2	Q54B19_DICDI	Q54b19_dicystostei	814	14	50.0	227	2	Q4QSP5_9LECA	Q4qsf5_pertuaria
742	14	50.0	214	2	Q5L6N1_CHLAB	Q5l6n1_chlamydomo	815	14	50.0	227	2	Q4QSL2_9LECA	Q4qsl2_pertuaria
743	14	50.0	214	2	Q9F6M0_RHJET	Q9f6m0_rhizobium_e	816	14	50.0	227	2	Q4QSK1_9LECA	Q4qsk1_pertuaria
744	14	50.0	214	2	Q53846_MYCTU	Q53846_mycobacteri	817	14	50.0	227	2	Q914U2_DROME	Q914u2_drosophila
745	14	50.0	214	2	Q7U158_MYCBO	Q7u158_mycobacteri	818	14	50.0	228	2	Q7UMT8_CABETL	Q7umt8_caenorhabdi
746	14	50.0	215	2	Q4WFS9_ASPFU	Q4wfs9_aspergillus	819	14	50.0	229	1	YAMB_RHISN	P55561_rhizobium_s
747	14	50.0	215	2	Q4PD16_USTMA	Q4pdl6_usfllago_ma	820	14	50.0	229	2	Q7Z972_9FUNG	Q7z972_fungal_endo
748	14	50.0	216	2	Q51CC4_EMENTI	Q51cc4_aspericella	821	14	50.0	229	2	Q7Z9U8_9FUNG	Q7z9u8_fungal_endo
749	14	50.0	216	2	Q5AQH5_EBENT	Q5aqh5_aspergillus	822	14	50.0	229	2	Q7Z9U9_9FUNG	Q7z9u9_fungal_endo
750	14	50.0	216	2	Q6KBI1_TREAVE	Q6kb19_tremetes_ve	823	14	50.0	229	2	Q7Z9K0_9FUNG	Q7z9k0_fungal_endo
751	14	50.0	216	2	Q4QSF2_9LECA	Q4qsf2_pertuaria	824	14	50.0	229	2	Q7Z9K1_9FUNG	Q7z9k1_fungal_endo
752	14	50.0	216	2	Q6ZM24_HUMAN	Q6zm24_homo_sapien	825	14	50.0	229	2	Q7Z9K2_9FUNG	Q7z9k2_fungal_endo
753	14	50.0	216	2	Q6NAG6_RHOPA	Q6nag6_rhodopseudo	826	14	50.0	229	2	Q7XB67_ORYSA	Q7xb67_oryza_sativ
754	14	50.0	217	2	Q6V276_9BACT	Q6v276_symbiont_ba	827	14	50.0	230	2	Q51UX3_MAGGR	Q51ux3_magnaporthe
755	14	50.0	217	2	Q5P3R3_AZOSUS	Q5p3r3_azaracus_sp	828	14	50.0	230	2	Q7Z970_9FUNG	Q7z970_fungal_endo
756	14	50.0	217	2	Q63PDS_BURPS	Q63pds_burkholderi	829	14	50.0	230	2	Q7Z971_9FUNG	Q7z971_fungal_endo
757	14	50.0	217	2	Q6ZEL4_BURMA	Q6zel4_burkholderi	830	14	50.0	231	2	Q8QSR7_9Z2Z2	Q8qsr7_unidentifie
758	14	50.0	217	2	Q8PKP6_XANAC	Q8pkp6_xanthomonas	831	14	50.0	232	2	Q27423_DROVI	Q27423_drosophila
759	14	50.0	217	2	Q8QSR6_9Z2Z2	Q8qsr6_unidentifie	832	14	50.0	233	2	Q4RE76_TETNG	Q4re76_tetradodon_n
760	14	50.0	217	2	Q38020_9VIRU	Q38020_porcato_viru	833	14	50.0	234	2	Q69HP3_CIOIN	Q69hp3_clona_inces
761	14	50.0	220	2	Q4QSH6_9LECA	Q4qsh6_pertuaria	834	14	50.0	234	2	Q9VIR4_DROME	Q9vir4_drosophila



835	14	50.0	234	2	Q98AP0_RHIL0	Q98AP0_rhizobium 1	908	14	50.0	270	2	Q46004_CABEL	Q46004_caenorhabdi
836	14	50.0	235	2	Q56PB1_9METZ	Q56PB1_uncultured	909	14	50.0	270	2	Q6EFC7_ACIAD	Q6EFC7_acinetobact
837	14	50.0	235	2	Q56PB3_9METZ	Q56PB3_uncultured	910	14	50.0	270	2	Q6JUBK1_ZEADI	Q6JUBK1_zea diploue
838	14	50.0	236	2	Q56P99_9METZ	Q56P99_uncultured	911	14	50.0	274	2	Q6TN86_METEPA	Q6TN86_methylobact
839	14	50.0	236	2	Q4SZK9_TETNG	Q4SZK9_tetradodon n	912	14	50.0	274	2	Q6PPN6_XYLEPA	Q6PPN6_xylella faa
840	14	50.0	239	2	Q9YB31_9ASCO	Q9YB31_phoma sp. c	913	14	50.0	274	2	Q87BC3_XYIPT	Q87BC3_xylella faa
841	14	50.0	239	2	Q5C403_SCHJA	Q5C403_schistosoma	914	14	50.0	274	2	Q8PLP2_XANAC	Q8PLP2_xanthomonas
842	14	50.0	239	2	Q7R376_GIALA	Q7R376_giardia lam	915	14	50.0	275	2	Q81109_WHEAT	Q81109_wheat
843	14	50.0	239	2	Q5BWM2_XANCA	Q5BWM2_xanthomonas	916	14	50.0	275	2	Q7W617_BORPA	Q7W617_bordelella
844	14	50.0	239	2	Q5BWM2_XANCA	Q5BWM2_xanthomonas	917	14	50.0	275	2	Q7W1F9_BORPA	Q7W1F9_bordelella
845	14	50.0	240	1	M810_AEATH	P92519_arabidopsis	918	14	50.0	275	2	Q7VYV2_BORPE	Q7VYV2_bordelella
846	14	50.0	240	2	Q5IEN0_9ARCH	Q5IEN0_uncultured	919	14	50.0	276	2	Q96GV4_HUMAN	Q96GV4_homo sapien
847	14	50.0	240	2	Q5IEN3_9ARCH	Q5IEN3_uncultured	920	14	50.0	276	2	Q9NBD0_GLOMR	Q9NBD0_gloeosira mo
848	14	50.0	240	2	Q9YB64_XANPA	Q9YB64_aspergillus	921	14	50.0	276	2	Q6UBK0_ZEADI	Q6UBK0_zea diploue
849	14	50.0	240	2	Q9HBS1_HUMAN	Q9HBS1_homo sapien	922	14	50.0	276	2	Q4UTS4_XANCP	Q4UTS4_xanthomonas
850	14	50.0	240	2	Q810D5_DROME	Q810D5_drosophila	923	14	50.0	276	2	Q5PEI1_SALPA	Q5PEI1_salmonella
851	14	50.0	241	2	Q9YB32_PENPA	Q9YB32_penicillium	924	14	50.0	276	2	Q8Z443_SALTI	Q8Z443_salmonella
852	14	50.0	241	2	Q9YB63_ASPPA	Q9YB63_aspergillus	925	14	50.0	276	2	Q8E9W2_XANCP	Q8E9W2_xanthomonas
853	14	50.0	242	2	Q5IEM9_9ARCH	Q5IEM9_uncultured	926	14	50.0	276	2	Q4S370_TETNG	Q4S370_tetradodon n
854	14	50.0	244	2	Q9Y100_9NROP	Q9Y100_ectodeamia	927	14	50.0	277	2	Q9VFN8_DROME	Q9VFN8_drosophila
855	14	50.0	244	2	Q4U052_XANCP	Q4U052_xanthomonas	928	14	50.0	278	2	Q8MOQ2_DROME	Q8MOQ2_drosophila
856	14	50.0	244	2	Q4F6V7_9BACT	Q4F6V7_uncultured	929	14	50.0	278	2	Q8MOQ2_DROME	Q8MOQ2_drosophila
857	14	50.0	244	2	Q8PHR5_XANAC	Q8PHR5_xanthomonas	930	14	50.0	278	2	Q9ZTA3_RHIME	Q9ZTA3_rhizobium m
858	14	50.0	245	2	Q6RKX5_GIBZE	Q6RKX5_gibberella	931	14	50.0	280	2	Q8YWG3_LOLMU	Q8YWG3_lolium mult
859	14	50.0	248	2	Q8PAJ7_XANCP	Q8PAJ7_xanthomonas	932	14	50.0	281	2	Q8UXM4_9VIRU	Q8UXM4_heliotais z
860	14	50.0	249	2	Q5AU21_EMENI	Q5AU21_aspergillus	933	14	50.0	282	2	Q7XL74_ORYSA	Q7XL74_oryza sativ
861	14	50.0	249	2	Q60N68_CABER	Q60N68_caenorhabdi	934	14	50.0	283	2	Q7D973_MYCTU	Q7D973_mycobacteri
862	14	50.0	249	2	Q16420_CABEL	Q16420_caenorhabdi	935	14	50.0	285	1	Q7R383_GIALA	Q7R383_giardia lam
863	14	50.0	250	1	BTG4_MOUSE	Q70552_mus musculu	936	14	50.0	286	1	CXA6_RAT	CXA6_rattus norv
864	14	50.0	250	2	Q61134_CABER	Q61134_caenorhabdi	937	14	50.0	286	2	Q8BRN2_MOUSE	Q8BRN2_mus musculu
865	14	50.0	250	2	Q16418_CABEL	Q16418_caenorhabdi	938	14	50.0	286	2	Q4RV58_TETNG	Q4RV58_tetradodon n
866	14	50.0	250	2	Q16419_CABEL	Q16419_caenorhabdi	939	14	50.0	287	2	Q8MWJ7_9ASCI	Q8MWJ7_boletenia v1
867	14	50.0	250	2	Q5SND6_ORYSA	Q5SND6_oryza sativ	940	14	50.0	288	2	Q4I937_GIBBE	Q4I937_gibberella
868	14	50.0	250	2	Q5FNU7_GILIOX	Q5FNU7_glycinebact	941	14	50.0	288	2	Q9XV81_CABEL	Q9XV81_caenorhabdi
869	14	50.0	250	2	Q5HWV9_CAMJR	Q5HWV9_campylobact	942	14	50.0	289	2	Q6H604_ORYSA	Q6H604_oryza sativ
870	14	50.0	250	2	Q925T9_MOUSE	Q925T9_mus musculu	943	14	50.0	289	2	Q7OV44_GIALA	Q7OV44_giardia lam
871	14	50.0	250	2	Q6NXX4_MOUSE	Q6NXX4_mus musculu	944	14	50.0	289	2	Q9VU78_DROME	Q9VU78_drosophila
872	14	50.0	251	2	Q4P139_USYMA	Q4P139_usciliago ma	945	14	50.0	289	2	Q4V9U3_BRAKE	Q4V9U3_brechydantio
873	14	50.0	253	2	Q4IACO_GIBZE	Q4IACO_gibberella	946	14	50.0	290	2	Q752M4_ASHGO	Q752M4_ashbya goss
874	14	50.0	253	2	Q6ZOX9_HUMAN	Q6ZOX9_homo sapien	947	14	50.0	290	2	Q9VX17_DROME	Q9VX17_drosophila
875	14	50.0	253	2	Q7UY70_RHOBA	Q7UY70_rhodopirell	948	14	50.0	290	2	Q6GHV3_BRAKE	Q6GHV3_brechydantio
876	14	50.0	254	2	Q6YUY0_ORYSA	Q6YUY0_oryza sativ	949	14	50.0	291	2	Q7XZP8_ORYSA	Q7XZP8_oryza sativ
877	14	50.0	254	2	Q55700_PRRSV	Q55700_porcine rep	950	14	50.0	291	2	Q6H5L8_ORYSA	Q6H5L8_oryza sativ
878	14	50.0	254	2	Q513H3_PRRSV	Q513H3_porcine rep	951	14	50.0	291	2	Q6ES25_ORYSA	Q6ES25_oryza sativ
879	14	50.0	254	2	Q6A521_PRRSV	Q6A521_porcine rep	952	14	50.0	292	2	Q8RIY7_MOUSE	Q8RIY7_mus musculu
880	14	50.0	254	2	Q84965_PRRSV	Q84965_porcine rep	953	14	50.0	293	2	Q7U3U8_SYNPX	Q7U3U8_synechococc
881	14	50.0	255	2	Q8N0W7_HUMAN	Q8N0W7_homo sapien	954	14	50.0	293	2	Q57124_9REOV	Q57124_coho salmon
882	14	50.0	255	2	Q56P90_9FUNG	Q56P90_fungal sp.	955	14	50.0	293	2	Q6ETW1_XENLA	Q6ETW1_xenopus lae
883	14	50.0	255	2	Q4Q1V5_LEIWA	Q4Q1V5_leishmania	956	14	50.0	296	2	Q4I1N2_GIBBE	Q4I1N2_gibberella
884	14	50.0	257	1	Q1095_HUMAN	Q1095_homo sapien	957	14	50.0	297	2	Q5ZNN2_9VIRU	Q5ZNN2_trypanosoma
885	14	50.0	257	2	Q5Z9G6_ORYSA	Q5Z9G6_oryza sativ	958	14	50.0	298	2	Q99159_TRYBU	Q99159_cotesia con
886	14	50.0	258	2	Q60N66_CABER	Q60N66_caenorhabdi	959	14	50.0	298	2	Q5VS69_ORYSA	Q5VS69_oryza sativ
887	14	50.0	258	2	Q9Y0B9_DROME	Q9Y0B9_drosophila	960	14	50.0	298	2	Q7U7N8_SYNPX	Q7U7N8_synechococc
888	14	50.0	258	2	Q00946_9HYMN	Q00946_tetrahymena	961	14	50.0	299	2	Q5R0R8_IDILO	Q5R0R8_idiomarina
889	14	50.0	258	2	Q5PAGO_ANAMM	Q5PAGO_anaplasma m	962	14	50.0	300	2	Q22513_SOYBN	Q22513_glycine max
890	14	50.0	258	2	Q8QVG2_ADBBA	Q8QVG2_bovine aden	963	14	50.0	300	2	Q4ZTF5_PSSBY	Q4ZTF5_pseudomonas
891	14	50.0	259	2	Q7S669_NEUCRA	Q7S669_neurospora	964	14	50.0	300	2	Q9YGD9_ONCMY	Q9YGD9_oncorynchus
892	14	50.0	259	2	Q83212_TREPA	Q83212_treponeema p	965	14	50.0	301	2	Q7V945_PROMM	Q7V945_prochlorococ
893	14	50.0	261	2	Q94EL3_SORHL	Q94EL3_sorghum hal	966	14	50.0	301	2	Q8BN70_MOUSE	Q8BN70_mus musculu
894	14	50.0	261	2	Q5BR83_RAT	Q5BR83_rattus norv	967	14	50.0	301	2	Q6ELZ9_BRAKE	Q6ELZ9_brechydantio
895	14	50.0	262	2	Q5DC20_SCHJA	Q5DC20_schistosoma	968	14	50.0	301	2	Q4VBS7_BRAKE	Q4VBS7_brechydantio
896	14	50.0	264	2	Q27253_CABEL	Q27253_mycohabdi	969	14	50.0	302	2	Q7X546_9ACTO	Q7X546_actinoplan
897	14	50.0	264	2	Q08343_MYCTU	Q08343_mycobacteri	970	14	50.0	302	2	Q4SZH4_TETNG	Q4SZH4_tetradodon n
898	14	50.0	264	2	Q7VFP02_MYCBO	Q7VFP02_mycobacteri	971	14	50.0	304	2	Q4LETA_9ILILI	Q4LETA_salicetaria
899	14	50.0	267	2	Q6S5T9_FUGRU	Q6S5T9_fugu rubrip	972	14	50.0	304	2	Q9CUT3_MOUSE	Q9CUT3_mus musculu
900	14	50.0	268	2	Q81110_WHEAT	Q81110_triticum ae	973	14	50.0	306	2	Q4UTZ8_XANCP	Q4UTZ8_xanthomonas
901	14	50.0	268	2	Q5H5W0_XANOR	Q5H5W0_xanthomonas	974	14	50.0	306	2	Q8EPD32_XANCP	Q8EPD32_xanthomonas
902	14	50.0	269	2	Q4Q5X9_LEIWA	Q4Q5X9_leishmania	975	14	50.0	307	2	Q94IDS_ORYSA	Q94IDS_oryza sativ
903	14	50.0	269	2	Q5GYN9_XANOR	Q5GYN9_xanthomonas	976	14	50.0	308	2	Q7R414_GIALA	Q7R414_giardia lam
904	14	50.0	269	2	Q6Z2Z22_BURPS	Q6Z2Z22_burkholderi	977	14	50.0	309	1	CF188_HUMAN	CF188_homo sapien
905	14	50.0	269	2	Q8KKG2_BURPS	Q8KKG2_burkholderi	978	14	50.0	309	1	CF188_MOUSE	CF188_mus musculu
906	14	50.0	269	2	Q8KKG3_BURPS	Q8KKG3_burkholderi	979	14	50.0	309	1	CF188_RAT	CF188_rattus norv
907	14	50.0	269	2	Q6Z2U6_BURPS	Q6Z2U6_burkholderi	980	14	50.0	309	2	Q9VZG9_DROME	Q9VZG9_drosophila



```

RESULT 2
O8K340 MOUSE PRELIMINARY; PRT; 347 AA.
ID O8K340;
AC O8K340;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, last annotation update)
DE 4632411J06Rik;
GN Name=4632411J06Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OC NCBI_Taxid=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CZECH II;
RC TISSUE=Mammary tumor metastasized to lung. MMTV-LTR/Mut1 model.
RX Expression driven by an MMTV-LTR enhancer.;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stalcenon M., Soares M.B., Ronaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield J.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CZECH II;
RC TISSUE=Mammary tumor metastasized to lung. MMTV-LTR/Mut1 model.
RC Expression driven by an MMTV-LTR enhancer.;
RA Strusberg R.L. Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028854; AAH28854.1; -; mRNA.
DR MGI; MGI:1925998; 4632411J06Rik.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000159; RA.
DR PROSITE; PS50200; RA; 1.
FT NON_TER 1
FT SEQUENCE 347 AA; 39161 MW; 306E40D38A14B3DE CRC64;

```

```

Query Match 64.3%; Score 18; DB 2; Length 347;
Best Local Similarity 16.7%; Pred. No. 5.1e-11;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

OY 1 CXXXXXXXXXXC 12
DB 135 C8STSSSTASSC 146

```

```

RESULT 3
O8COD3 MOUSE PRELIMINARY; PRT; 362 AA.
ID O8COD3;
AC O8COD3;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, last annotation update)
DE Mus musculus 13 days embryo male testis cDNA, RIKEN full-length

```

```

DE enriched library, clone:6030474M07 product:hypothetical serine-rich
DE region/Aminoacyl-transfer RNA synthetases class-II containing protein,
DE full insert sequence.
GN Name=4632411J06Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OC NCBI_Taxid=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RT Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konho H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kaibawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giusti C., King B., Koehli H.,
RA Kuehl P., Lewis S., Matsumoto Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Oikio T., Furuno M., Aono H., Balarelli R., Barsh G.,
RA Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guestenrich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombere P.,
RA Morone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RT Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RT Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe K.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsushita S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RT Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.

```

RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishii K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,  
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AK031660; BAC27498.1; -; mRNA.  
 DR MGI; MGI:1925998; 4632411J06Rik.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR000159; RA.  
 DR SMART; SM00314; RA; 1.  
 DR PROSITE; PS50200; RA; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 362 AA; 40934 MW; 32B1047AC07B987 CRC64;

Query Match 64.3%; Score 18; DB 2; Length 362;  
 Best Local Similarity 16.7%; Pred. No. 5.2e-11;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
 DB 199 C8STSSSTSSC 210

RESULT 4  
 O6AV38 ORYSA PRELIMINARY; PRT; 367 AA.  
 AC O6AV38;  
 DT 25-OCT-2004 (TRENBLrel. 28, Created)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)  
 DE Putative Cyclin.  
 GN Name=OSJNBA0063J18.4;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,  
 RA Overton II L.L., Taiter T., Kim M.M., Bera J.J., Jin S.S.,  
 RA Padrosh D.W., Tallon L.J., Koo H., Ziemann V., Hsiao J., Blunt S.,  
 RA Vanden S.S., Riedmuller S.B., Peterback T.T., Feldlyum T.V.,  
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,  
 RA White O., Salzberg S.L., Fraser C.M.;  
 RT "Oryza sativa chromosome 3 BAC OSJNBA0063J18 genomic sequence."  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Buell R.;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 CC -; SIMILARITY: Belongs to the cyclin family.  
 DR EMBL; AC107206; AAT77041.1; -; Genomic\_DNA.  
 DR Gramene; O6AV38; -;  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0000734; P:regulation of cell cycle; IEA.  
 DR InterPro; IPR006670; Cyclin.  
 DR InterPro; IPR004367; Cyclin\_Cterm.  
 DR InterPro; IPR006671; Cyclin\_M.  
 DR Pfam; PF02984; Cyclin\_C; 1.  
 DR Pfam; PF00134; Cyclin\_N; 1.  
 DR SMART; SM00385; CYCLIN; 1.  
 KW Cyclin.  
 SQ SEQUENCE 367 AA; 38727 MW; 023BF8C0CAD1D79 CRC64;

Query Match 64.3%; Score 18; DB 2; Length 367;  
 Best Local Similarity 16.7%; Pred. No. 5.2e-11;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
 DB 75 C8STSSSTSSC 86

RESULT 5  
 O8BL43 MOUSE PRELIMINARY; PRT; 584 AA.  
 ID O8BL43;  
 AC O8BL43;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length  
 DE enriched library, clone:B230384H21 product:hypothetical Serine-rich  
 DE region/Aminoacyl-transfer RNA synthetases class-II containing protein,  
 DE full insert sequence. (fragment).  
 GN Name=4632411J06Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RT Carninci P., Hayashizaki Y.;  
 RL "High-efficiency full-length cDNA cloning."  
 RL Meth. Enzymol. 303:19-44 (1999).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guscinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarte P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weiss C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690 (2001).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RA STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RA The PANTOM Consortium.  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573 (2002).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE.  
 RA STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=20530913; PubMed=1076861; DOI=10.1101/9r.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,  
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Matshiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Kato H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saeki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akanita S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK046433; BAC32726.1; -; mRNA.  
DR MGI: MGI:1925998; 4632411J06Rik.  
DR GO: GO:0007165; P:signal transduction; IEA.  
DR InterPro: IPR000159; RA.  
DR SMART: SM00314; RA; 1.  
DR PROSITE: PS50200; RA; 1.  
KM Hypothetical protein.  
FT NON TER 1  
SQ SEQUENCE 584 AA; 65345 MW; 941D03439114CC6E CRC64;  
Query Match 64.3%; Score 18; DB 2; Length 584;  
Best Local Similarity 16.7%; Pred. No. 5.9e-11;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 1 CXXXXXXXXXXC 12  
Db 275 CSSTSSSTASC 286  
RESULT 6  
ID Q8LP13\_PEA PRELIMINARY; PRT; 730 AA.  
AC Q8LP13;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ABA insensitive 3.  
GN Name=abi3;  
OS Pisum sativum (Garden pea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciales; Pisum.  
OX NCBI\_TaxID=3888;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Axillary bud;  
RA Nakako A., Mori H.,  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB080195; BAC10553.1; -; mRNA.  
DR GO: GO:0003677; F:DNA binding; IEA.  
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro: IPR003340; TF\_B3.  
DR Pfam: PF02362; B3; 1.

DR PROSITE: PS50863; B3; 1.  
SQ SEQUENCE 730 AA; 80868 MW; 0A5BE9A7350AB8B2 CRC64;  
Query Match 64.3%; Score 18; DB 2; Length 730;  
Best Local Similarity 16.7%; Pred. No. 6.2e-11;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 1 CXXXXXXXXXXC 12  
Db 55 CSSSSSSASSSC 66  
RESULT 7  
ID Q9L012\_ARATH PRELIMINARY; PRT; 35 AA.  
AC Q9L012;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE F1504.2.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopses.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Ecker J.R.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Cheuk R., Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
RA Khan S., Kim C., Alfieri H., Bei B., Chin C., Chou J., Choi E.,  
RA Com L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
RA Lee J., Lenz C., Li U., Liu A., Liu J., Liu S., Mukharzky N.,  
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
RA Thayer A., Toriumi M., Vaynsberg M., Yu G., Davis R., Federspiel N.,  
RA Theologis A., Ecker J.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC007887; AAF79343.1; -; Genomic DNA.  
SQ SEQUENCE 35 AA; 3663 MW; CB4370A9D6A5D1F4 CRC64;  
Query Match 60.7%; Score 17; DB 2; Length 35;  
Best Local Similarity 16.7%; Pred. No. 1.9e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 1 CXXXXXXXXXXC 12  
Db 15 CSIASTAAATSC 26  
RESULT 8  
ID Q24960\_GIALA PRELIMINARY; PRT; 137 AA.  
AC Q24960;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Trophozoite surface protein II (Fragment).  
GN Name=TSPI1/L2;  
OS Giardia lamblia (Giardia intestinalis).  
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
OX NCBI\_TaxID=5741;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC MEDLINE=98096860; PubMed=9435134;  
RA Ey P.L., Mansouri M., Kulda J., Nohynkova E., Monis P.T.,  
RA Andrews R.H., Mayrhofer G.,  
RT "Genetic analysis of Giardia from hoofed farm animals reveals  
RT antiodontal-specific and potentially zoonotic genotypes.";  
RL J. Eukaryot. Microbiol. 44:626-635(1997).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.

```
RA Ey P.L., Mansouri M., Kulda J., Nohynkova E., Monis P.T.,  
RA Andrews R.H., Mayrhofer G.,  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U47631; AAB05399.1; -, Genomic_DNA.  
DR InterPro; IPR006209; EGF_1like.  
DR InterPro; IPR005127; Giardia_vsp.  
DR Pfam; PF03302; VSP; 1.  
DR PROSITE; PS01186; EGF_2; 1.  
FT NON_TER 1  
FT NON_TER 137  
SQ SEQUENCE 137 AA; 13875 MW; 42346969B5ECF37D CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 137;  
Best Local Similarity 16.7%; Pred. No. 2.7e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 22 CASATARAATC 33  
  
RESULT 9  
Q24987 GIALA PRELIMINARY; PRT; 137 AA.  
AC Q24987;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Surface protein (Fragment).  
GN Name: vsp417-4;  
OS Giardia lamblia (Giardia intestinalis).  
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
OX NCBI_Taxid=5741;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Ey P.L., Darby J.M., Mayrhofer G.,  
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L16973; AAB59198.1; -, Genomic_DNA.  
DR InterPro; IPR005127; Giardia_vsp.  
DR Pfam; PF03302; VSP; 1.  
FT NON_TER 1  
FT NON_TER 137  
SQ SEQUENCE 137 AA; 14148 MW; A340072153005D58 CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 137;  
Best Local Similarity 16.7%; Pred. No. 2.7e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 22 CTSTARTATATC 33  
  
RESULT 10  
Q7XHS1 ORYSA PRELIMINARY; PRT; 210 AA.  
AC Q7XHS1;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE 2Fe-2S iron-sulfur cluster protein-like.  
GN Name: P0477A12.11;  
OS Oryza sativa (japonica cultivar-group). Embryophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Eukaryophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhacridae; Oryzaceae; Oryza.  
OX NCBI_Taxid=39947;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Sasaki T., Matsumoto T., Katayose Y.,  
RT "Oryza sativa japonica (GAS) genomic DNA, chromosome 7, PAC  
RT clone: P0477A12.";  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AF005190; BAC80058.1; -, Genomic_DNA.  
DR Gramene; Q7XHS1; -;  
DR GO; GO:0005489; Electron transporter activity; IEA.  
DR GO; GO:0005506; P:iron ion binding; IEA.  
DR GO; GO:0046872; P:metal ion binding; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR01041; Ferredoxin.  
DR InterPro; IPR012675; Ferredoxin_fold.  
DR Pfam; PF00111; Fer2; 1.  
SQ SEQUENCE 210 AA; 22629 MW; 6BCDC4F408B21E1E CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 210;  
Best Local Similarity 16.7%; Pred. No. 3e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 25 CAVATATTTTSSC 36  
  
RESULT 11  
Q9XY90 GIALA PRELIMINARY; PRT; 234 AA.  
AC Q9XY90;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Variant-specific surface protein (Fragment).  
GN Name=vsp417-4;  
OS Giardia lamblia (Giardia intestinalis).  
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
OX NCBI_Taxid=5741;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Ad-2;  
RX MEDLINE=99229863; PubMed=10215024; DOI=10.1016/S0166-6851(98)00183-2;  
RA Ey P.L., Darby J.M., Mayrhofer G.,  
RT "A new locus (vsp417-4) belonging to the tsa417-like subfamily of  
RT variant-specific surface protein genes in Giardia intestinalis.";  
RL Mol. Biochem. Parasitol. 99:55-68(1999).  
DR EMBL; AF065600; AAD28789.1; -, Genomic_DNA.  
DR InterPro; IPR006212; Furin_repeat.  
DR InterPro; IPR005127; Giardia_vsp.  
DR InterPro; IPR006210; IBBF.  
DR Pfam; PF03302; VSP; 1.  
DR SMART; SM00181; EGF; 2.  
DR SMART; SM00261; FU; 1.  
FT NON_TER 1  
FT NON_TER 234  
SQ SEQUENCE 234 AA; 23565 MW; FDC75E280AF7D517 CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 234;  
Best Local Similarity 16.7%; Pred. No. 3.1e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 22 CTSATARAATC 33  
  
RESULT 12  
O18039 CAEEL PRELIMINARY; PRT; 245 AA.  
AC O18039;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical protein T06C12.14.  
GN ORFNames=T06C12.14;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoa; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI_Taxid=6239;
```

RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=Briscot N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RG The C. elegans sequencing consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology.";  
RL Science 282:2012-2018 (1998).  
DR EMBL: Z81116; CAB03300.1; -; Genomic\_DNA.  
DR PIR: T24565; T24565.  
DR Ensembl: T06C12.14; Caenorhabditis elegans.  
DR Wormbase: WBGene00011521; T06C12.14.  
DR Wormpep: T06C12.14; CE16364.  
DR InterPro: IPR003582; SHKT.  
DR Pfam: PF01549; SHKT; 3.  
DR SMART: SM00254; SHKT; 4.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 245 AA; 25713 MW; BDE40649C966769A CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 245;  
Best Local Similarity 16.7%; Pred. No. 3.1e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 193 CASATTTTSSCTC 204  
  
RESULT 13  
Q5WQW8 CRYNE PRELIMINARY; PRT; 306 AA.  
ID Q5WQW8; CRYNE PRELIMINARY;  
AC Q5WQW8;  
DT 25-OCT-2004 (TREMBlrel. 28, Created)  
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
DE Sec53p.  
GN Name=SEC53;  
OS Cryptococcus neoformans var. neoformans.  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.  
OX NCBI\_TaxID=40410;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.  
RA Jambon G.;  
RT "SEC53 encodes a putative phosphomannomutase.";  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY426175; AAR84595.1; -; Genomic\_DNA.  
DR GO: GO:0005737; Cytoplasm; IEA.  
DR GO: GO:0003824; Catalytic activity; IEA.  
DR GO: GO:0004615; Phosphomannomutase activity; IEA.  
DR GO: GO:0019307; P-mannose biosynthesis; IEA.  
DR GO: GO:0008152; P-metabolism; IEA.  
DR InterPro: IPR006379; HAD\_SF\_IIB.  
DR Pfam: PF03332; PMM; 1.  
DR TIGRFAMs: TIGR01484; HAD-SF-IIB; 1.  
SQ SEQUENCE 306 AA; 33844 MW; FB9DB2CF1A3EDDEA CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 306;  
Best Local Similarity 16.7%; Pred. No. 3.3e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 5 CSASATTTATTC 16  
  
RESULT 14  
Q6ASQ2 ORISA PRELIMINARY; PRT; 378 AA.  
ID Q6ASQ2; ORISA PRELIMINARY;  
AC Q6ASQ2;  
DT 25-OCT-2004 (TREMBlrel. 28, Created)  
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
DE Hypothetical protein OSUNBa007J22.21.  
GN Name=OSUNBa007J22.21;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,  
RA Chen Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,  
RA Chao Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiiao S.-H.,  
RA Heung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,  
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,  
RA Wu H.-P., Shaw J.-F.;  
RT "Oryza sativa BAC OSUNBa007J22 genomic sequence.";  
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC148814; AAT93980.1; -; Genomic\_DNA.  
DR Gramene; Q6ASQ2; -.  
KW Hypothetical protein.  
SQ SEQUENCE 378 AA; 37703 MW; 342586481D07BF5B CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 378;  
Best Local Similarity 16.7%; Pred. No. 3.5e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 CXXXXXXXXXXC 12  
DB 180 CVAAAAAATPAC 191  
  
RESULT 15  
QALGT2\_9BURK PRELIMINARY; PRT; 396 AA.  
ID QALGT2\_9BURK;  
AC QALGT2;  
DT 13-SEP-2005 (TREMBlrel. 31, Created)  
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)  
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=Bcen242DRAFT\_0168;  
OS Burkholderia cenocepacia H12424.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.  
OX NCBI\_TaxID=331272;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA STRAIN=H12424;  
RC US DOE Joint Genome Institute (JGI-PGF);  
RG US DOE Joint Genome Institute (JGI-PGF);  
RA Copeland A., Lucas S., Lapidus A., Barry K., Deter C., Glavina T.,  
RA Hammon N., Israni S., Pluck S., Richardson P.;  
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia  
H12424.";  
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=H12424;  
RG US DOE Joint Genome Institute (JGI-ORNL);  
RA Larter P., Land M.;  
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia  
H12424.";  
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
preliminary data.  
KW Hypothetical protein.  
DR EMBL: AAL01000139; EAM15270.1; -; Genomic\_DNA.  
SQ SEQUENCE 396 AA; 41189 MW; C829063DC312B057 CRC64;  
  
Query Match 60.7%; Score 17; DB 2; Length 396;  
Best Local Similarity 16.7%; Pred. No. 3.5e-09;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;



```

QY      1 CXXXXXXXXXXC 12
      |
DB      18 CSSSDAAAAAAC 29

RESULT 16
O8C8X1_MOUSE PRELIMINARY;      PRT;      429 AA.
ID O8C8X1_MOUSE PRELIMINARY;
AC O8C8X1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult retina cDNA, RIKEN full-length enriched library,
DE clone: A330006D11 product: hypothetical protein, full insert
DE sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
(2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai U., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Komno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsumoto H.A., Ashburner M., Batalov S., Casavant T.,
RA Pleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbetsu P.,
RA Norcone P., Ring B., Ringwald M., Rodriguez I., Sakemoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
(3)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RA The PANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
(4)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
(5)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

```

```

RA Kono H., Akiyama J., Nishi K., Katsunori T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai U.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multichannel sequencer.";
RL Genome Res. 10:1757-1771(2000).
(6)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RA Adachi J., Aikawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa H., Kasukawa T.,
RA Katoh H., Kawai U., Kojima Y., Kondo S., Komno H., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akehira S., Takada Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK044299; BAC11860.1; -; mRNA.
DR Ensembl: ENSMUSG0000014198; Mus musculus.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR InterPro: IPR007087; Znf_C2H2.
DR InterPro: IPR003604; Znf_U1.
DR SMART: SM00355; ZNF_C2H2_3.
DR SMART: SM00451; ZNF_U1; 3.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 429 AA; 44564 MW; 211C9F431F8F7FEF CRC64;

Query Match      60.7%; Score 17; DB 2; Length 429;
Best Local Similarity 16.7%; Pred. No. 3,6e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      1 CXXXXXXXXXXC 12
      |
DB      181 CDAAASSSSSSC 192

RESULT 17
O93H33_STRAW PRELIMINARY;      PRT;      940 AA.
ID O93H33_STRAW PRELIMINARY;
AC O93H33;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Transport system protein (Putative oligopeptide ABC transporter
DE substrate-binding protein).
GN Name=oppA2; OrderedLocustNames=SAV1149;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomyces; Streptomyces.
OC NCBI_TaxID=33903;
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.21143198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakai Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
(2)
RP NUCLEOTIDE SEQUENCE.

```



```

RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=2268306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AB070952; BAB69358.1; -; Genomic DNA.
DR EMBL; BA000030; BAC70860.1; -; Genomic DNA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000914; SBP_bac_5.
DR Pfam; PF00496; SBP_bac_5; 1.
SQ Complete proteome.
Query Match 540 AA; 58574 MW; 0052FEAD151B442C CRC64;
Best Local Similarity 60.7%; Score 17; DB 2; Length 540;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 12 CAAAAAATAATC 23

RESULT 18
ID Q9XTK3_GIALA PRELIMINARY; PRT; 667 AA.
AC Q9XTK3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE VSP417-3/A-II.
GN Name=vsp417-3/A-II;
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BRIS/83/HEPU/136;
RX MEDLINE=99053029; PubMed=9836309;
RA Ey P.L., Darby J.M., Mayrhofer G.;
RT "Comparison of tsaa117-like variant-specific surface protein (VSP)
RT genes in Giardia intestinalis and identification of a novel locus in
RT genetic Group II isolates.";
RL Parasitology 117:445-455(1998).
DR EMBL; AF033584; AAD03497.1; -; Genomic_DNA.
DR HSSP; O16119; IEZG.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR005127; Giardia_VSP.
DR InterPro; IPR006210; IEGR.
DR Pfam; PF03302; VSP; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00261; FU; 4.
SQ SEQUENCE 667 AA; 69123 MW; 77C64CF959441C0C CRC64;

Query Match 60.7%; Score 17; DB 2; Length 667;
Best Local Similarity 16.7%; Pred. No. 4e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 448 CTSTARTATATC 459

```

```

GN ORFNames=AN4907.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=227321;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nuebaum C., Abouelleil A., Allen N., Anderson S.,
RA Archchi H.M., Barna N., Bastien V., Bloom T., Boguslavsky L.,
RA Boukhalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Cornu B., DeArliano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hago B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kanat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., McLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Seery C., Rogov P.,
RA Roman J., Schauer S., Schuback R., Seaman S., Sivey P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of Aspergillus nidulans.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD0100084; BAA60985.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1031 AA; 113626 MW; 6616443769CF2EF CRC64;

Query Match 60.7%; Score 17; DB 2; Length 1031;
Best Local Similarity 16.7%; Pred. No. 4.5e-09;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 743 CTSSTSSASSTIC 754

RESULT 20
ID MUC5B_HUMAN STANDARD; PRT; 5703 AA.
AC Q9H8C4; Q00447; Q00573; Q14985; Q15494; Q95251; Q95451; Q14881;
AC Q99552; Q9DE28;
DT 28-FEB-2003 (Ref. 41, Created)
DT 28-FEB-2003 (Ref. 41, last sequence update)
DT 13-FEB-2005 (Ref. 48, last annotation update)
DE Mucin-5B precursor (Mucin 5 subtype B, tracheobronchial) (High
DE molecular weight salivary mucin M61) (Sublingual gland mucin).
GN Name=MUC5B; Synonyms=MUC5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE OF 1-1594.
RC Chen Y., Di Y.P., Wu R.;
RT "Molecular cloning of the amino-terminal and 5'-flanking region of the
RT human MUC5B mucin gene.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE OF 1-1325.
RX MEDLINE=99009274; PubMed=9790959; DOI=10.1006/dbrc.1998.9469;

```



```

FT CARBOHYD 5604 5604 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 5618 5618 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 5662 5662 N-linked (GlcNAc...) (Potential)
FT DISULFID 5594 5594 By similarity.
FT DISULFID 5622 5622 By similarity.
FT DISULFID 5626 5626 By similarity.
FT DISULFID 5643 5643 By similarity.
FT DISULFID 5645 5645 By similarity.
FT VARIANT 5137 5137 T->S (in dbSNP:2672788).
FT /FTID=VAR_014123.
FT CONFLICT 34 34 G->E (in Ref. 2).
FT CONFLICT 95 100 PPELGN->LPCLCK (in Ref. 2).
FT CONFLICT 104 104 S->C (in Ref. 2).
FT CONFLICT 142 142 E->K (in Ref. 1).
FT CONFLICT 225 225 R->S (in Ref. 2).
FT CONFLICT 330 331 PL->T (in Ref. 2).
FT CONFLICT 337 337 E->N (in Ref. 2).
FT CONFLICT 356 356 E->K (in Ref. 2).
FT CONFLICT 362 362 G->R (in Ref. 2).
FT CONFLICT 369 369 Missing (in Ref. 2 and 3).
FT CONFLICT 374 374 D->N (in Ref. 2).
FT CONFLICT 393 394 RT->TR (in Ref. 2).
FT CONFLICT 468 469 RK->GR (in Ref. 2).
FT CONFLICT 512 512 L->P (in Ref. 2).
FT CONFLICT 585 587 GAA->AH (in Ref. 3).
FT CONFLICT 601 601 A->S (in Ref. 3).
FT CONFLICT 628 629 DP->RS (in Ref. 2).
FT CONFLICT 633 633 F->L (in Ref. 2).
FT CONFLICT 676 676 A->P (in Ref. 3).
FT CONFLICT 701 701 R->P (in Ref. 3).
FT CONFLICT 752 752 E->K (in Ref. 2).

```

Query Match Score 17; DB 1; Length 5703;

Best Local Similarity 16.7%; Pred. No. 1.6e-09; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 5062 CTSASATMAARC 5073

```

RESULT 21
O6EUA6 ORYSA PRELIMINARY; PRT; 93 AA.
AC O6EUA6.
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DE Hypothetical protein OJ116_C12.20.
GN Name=OJ116_C12.20;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCB1_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004134; BAD27763.1; -; Genomic_DNA.
DR Gramene; O6EUA6; -.
KW Hypothetical protein.
SQ SEQUENCE 93 AA; 9912 MW; 10814C0A5B0F123 CRC64;

```

Query Match Score 16; DB 2; Length 93;

Best Local Similarity 16.7%; Pred. No. 1.6e-07; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 14 CTSASAGDATSC 25

```

RESULT 22
O850L9 ORYSA PRELIMINARY; PRT; 100 AA.
AC O850L9.
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative topoisomerase.
OS Name=OSJNBA0032E21.11;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCB1_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Eastman A.P., Smith S.C., Bertin N., Liang C., Najjar F.Z., Pratt L.H.,
RA Cordomier-Pratt M.-M.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF377947; AAO32310.1; -; Genomic_DNA.
DR Gramene; O850L9; -.
DR GO; GO:0016853; P:isomerase activity; IEA.
KW Isomerase.
SQ SEQUENCE 100 AA; 11026 MW; FA5BFLC9B0965999 CRC64;

```

Query Match Score 16; DB 2; Length 100;

Best Local Similarity 16.7%; Pred. No. 1.6e-07; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 60 CSSASASMAASP 71

```

RESULT 23
O6ZUJ1 BURMA PRELIMINARY; PRT; 100 AA.
AC O6ZUJ1.
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedlocusNames=BMA1578;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCB1_TaxID=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.040306101;
RA Nieman W.C., Deshazer D., Kim H.S., Tetteijn H., Nelson K.E.,
RA Feldlyum T.V., Ulrich R.L., Roming C.M., Brinkac L.M.,
RA Daugherty S.C., Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Kadane D., Romero C.M.,
RA Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome."
RT Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR EMBL; CP000010; AAU47930.1; -; Genomic_DNA.
DR TIGR; BMA1578; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 100 AA; 11000 MW; 89664E4C442D9F5 CRC64;

```

Query Match Score 16; DB 2; Length 100;

Best Local Similarity 16.7%; Pred. No. 1.6e-07; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12

DB 20 CTAATAATTATC 31

```

RESULT 24
CBPI_AJECA STANDARD; PRT; 110 AA.
ID CBPI_AJECA
AC 042720;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Calcium-binding protein precursor (CBP).
GN Name=CBP1;
OS Ajellomyces capsulata (Histoplasma capsulatum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Oxyenales; Ajellomycetaceae; Ajellomyces.
OX NCBI_TaxID=5037;

[1]
RP NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RC STRAIN=G186AR;
RX MEDLINE=98149309; PubMed=9489665;
RA Batanghari J.W., Deepe G.S. Jr., Di Cera E., Goldman W.E.;
RT "Histoplasma acquisition of calcium and expression of CBP1 during
RT intracellular parasitism.";
RL Mol. Microbiol. 27:531-539(1998).
[2]
RP CHARACTERIZATION, AND MASS SPECTROMETRY.
RC STRAIN=G186AR;
RX MEDLINE=98053973; PubMed=9393824;
RA Batanghari J.W., Goldman W.E.;
RT "Calcium dependence and binding in cultures of Histoplasma
RT capsulatum.";
RL Infect. Immun. 65:5257-5261(1997).
CC -!- FUNCTION: Involved in calcium binding and uptake in yeast phase.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- DEVELOPMENTAL STAGE: Expressed in the yeast phase during
CC infection, and not in mycelia.
CC -!- PTM: May be glycosylated.
CC -!- MASS SPECTROMETRY: MW=7858.0; METHOD=Electrospray; RANGE=33-110;
CC NOE=Ref.2.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF006209; AAC39354.1; -; Genomic_DNA.
CC KW Calcium; Direct protein sequencing; Glycoprotein; Signal.
CC FT SIGNAL 1 19 Potential.
CC FT CHAIN 1 110 Calcium-binding protein.
CC FT DISULFID 60 93 Potential.
CC FT DISULFID 65 76 Potential.
CC FT DISULFID 105 108 Potential.
CC SQ SEQUENCE 110 AA; 11014 MW; BD98F2FA1B4C77CE CRC64;

Query Match 57.1%; Score 16; DB 1; Length 110;
Best Local Similarity 16.7%; Pred. No. 1.7e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 65 CUSLSASSAAC 76

```

```

OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Han Z.;
RT Submitted (MAR-2005) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY810427; AAX26316.1; -; mRNA.
DR InterPro; IPR009081; ACP-like.
DR InterPro; IPR011899; ARM-like.
DR InterPro; IPR011992; EF-Hand_type.
KW Hypothetical protein.
SQ SEQUENCE 144 AA; 16762 MW; FF652713F976647B CRC64;

Query Match 57.1%; Score 16; DB 2; Length 144;
Best Local Similarity 16.7%; Pred. No. 1.8e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 15 CSSSYAVSTSC 26

RESULT 26
O9YBV0_AERPE PRELIMINARY; PRT; 152 AA.
ID O9YBV0_AERPE
AC O9YBV0;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein APL1499.
DE OrderedLocNames=APL1499;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcates;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KL;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyana A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix KL.";
RL DNA Res. 6:83-101(1999).
DR EMBL; BA000002; BAA80498.1; -; Genomic_DNA.
DR PIR; D72630; D72630.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 152 AA; 16241 MW; 4386978A0B6151BF CRC64;

Query Match 57.1%; Score 16; DB 2; Length 152;
Best Local Similarity 16.7%; Pred. No. 1.8e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 137 CSSTATTLSTRC 148

RESULT 27
O75LM9_ORYSA PRELIMINARY; PRT; 161 AA.
ID O75LM9_ORYSA
AC O75LM9;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative DNA topoisomerase (with alternative splicing).
GN Name=OSJNBa0047E24.8;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriarctoidae; Oryzaceae; Oryza.  
 OX NCBI\_Taxid=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,  
 RA Overton II L.L., Taitlin T., Kim M.M., Bera J.J., Jin S.S.,  
 RA Padgug D.W., Tallon L.J., Koo H., Ziemann V., Hsieh J., Blunt S.,  
 RA Vanaken S.S., Riedmuller S.B., Uteback T.T., Feldlyum T.V.,  
 RA Yang Q.Q., Haas B.J., Sub B.B., Peterson J.J., Quackenbush J.,  
 RA White O., Salzberg S.L., Fraser C.M.,  
 RA Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Buell R.;  
 RL Submitted (Jan-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC092556; AAR87258.1; -; Genomic\_DNA.  
 DR Gramene; Q75LM9; -;  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.  
 DR GO; GO:0016853; F:isomerase activity; IEA.  
 DR GO; GO:0006265; P:DNA topological change; IEA.  
 DR InterPro; IPR002815; DNA\_topII.  
 DR InterPro; IPR011991; Wing\_hlx\_DNA\_bd.  
 DR Pfam; PF04406; TPGA\_N.1.  
 DR PRINTS; PRO1550; TOF6AFAMILY.  
 DR DNA-binding; Isomerase; Transcription.  
 KW SEQUENCE 161 AA; 18051 MW; 41B12D2B71B76C3 CRC64;  
 SQ  
 Query Match 57.1%; Score 16; DB 2; Length 161;  
 Best Local Similarity 16.7%; Pred. No. 1.9e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXXXC 12  
 DB 60 CSSASAAAAPC 71

RESULT 28  
 Q7PY44\_ANOGA PRELIMINARY; PRT; 172 AA.  
 ID Q7PY44;  
 AC Q7PY44;  
 DT 01-MAR-2004 (TREMBlrel. 26, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE ENSANGP00000012250 (Fragment).  
 GN ORFNames=ENSANGG00000009761;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematoceera; Culicoidae; Culicidae;  
 OC Anophelinae; Anopheles.  
 OX NCBI\_Taxid=180454;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA STRAIN=PEST;  
 RA Anopheles genome sequencing Consortium;  
 RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAA80100887; EAA01398.1; -; Genomic\_DNA.  
 DR InterPro; IPR001092; HLH\_basic.  
 DR Pfam; PF00010; HLH; 1.  
 DR PROSITE; PS50888; HLH; 1.  
 FT NON\_TER 1  
 SEQUENCE 172 AA; 18464 MW; E857469A135A1884 CRC64;  
 Query Match 57.1%; Score 16; DB 2; Length 172;  
 Best Local Similarity 16.7%; Pred. No. 1.9e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
 DB 77 CXXXXSSSAGSEC 88  
 RESULT 29  
 Q6ZD90\_ORYSA PRELIMINARY; PRT; 176 AA.  
 ID Q6ZD90\_ORYSA  
 AC Q6ZD90;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Hypothetical protein P0438H08.27.  
 GN Name=P0438H08.27;  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriarctoidae; Oryzaceae; Oryza.  
 OX NCBI\_Taxid=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RL Submitted (Dec-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
 DR EMBL; AP004460; BAC99511.1; -; Genomic\_DNA.  
 DR Gramene; Q6ZD90; -;  
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.  
 DR GO; GO:0046872; F:metal ion binding; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0016567; P:protein ubiquitination; IEA.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4\_1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2, 1.  
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.  
 KW SEQUENCE 176 AA; 18608 MW; 79FE9D42C740112 CRC64;  
 SQ  
 Query Match 57.1%; Score 16; DB 2; Length 176;  
 Best Local Similarity 16.7%; Pred. No. 1.9e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 CXXXXXXXXXXC 12  
 DB 96 CRRAAAAAASC 107

RESULT 30  
 Q55014\_DICDI PRELIMINARY; PRT; 187 AA.  
 ID Q55014\_DICDI  
 AC Q55014;  
 DT 13-SEP-2005 (TREMBlrel. 31, Created)  
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=DPB0217903;  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 OX NCBI\_Taxid=44689;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA STRAIN=AX4;  
 RA Eichinger L., Pachepat J.A., Gloeckner G., Rajandream M.-A.,  
 RA Sugang R., Berrihan M., Song J., Olesch R., Szafarski K., Xu Q.,  
 RA Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,  
 RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,  
 RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,  
 RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,  
 RA Farbrother P., Desany B., Just E., Morio T., Rost R., Church C.,  
 RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,  
 RA Muzny D., Moutier T., Pain A., Lu M., Harper D., Lunday R.,  
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,  
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,

RA Louised H., Mungall K., Oliver K., Price C., Quail M.A.,  
 RA Urushihara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,  
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivy A.,  
 RA Sugano S., White B., Walker D., Woodward J., Winkler T., Tanaka Y.,  
 RA Shalensky G., Gishler R., Weinstock G., Rosenthal A., Cox E.C.,  
 RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,  
 RA Williams V., Dear P.H., Noegel A.A., Barrett B., Kuapa A.,  
 RT "The genome of the social amoeba *Dictyostelium discoideum*.";  
 RT Nature 0:0-0(2005).  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL:AAFI01000042; EMBL6047.1; -; Genomic\_DNA.  
 DR Hypothetical protein.  
 KW Hypothetical protein.  
 SQ SEQUENCE 187 AA; 20124 MW; 8D78DE804DFBC104 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 187;  
 Best Local Similarity 16.7%; Pred. No. 1.9e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12  
 Db 65 CYSYSSSSSFC 76

RESULT 31  
 082LKO\_STRAW PRELIMINARY; PRT; 208 AA.  
 ID 082LKO;  
 AC 082LKO;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN Ordered locus names=SAV2010;  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycetaceae; Streptomycetaceae; Streptomycetes.  
 OC NCBI\_Taxid=33903;  
 RX NCBI\_Taxid=33903;  
 RA NUCLEOTIDE SEQUENCE.

RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RC MEDLINE=21477403; PubMed=12692562; DOI=10.1038/nbt820;  
 RX OMURA S., IKEDA H., ISHIKAWA J., HANAMOTO A., TAKAHASHI C.,  
 RA SHIROSE M., TAKAHASHI Y., HORIKAWA H., NAKAZAWA H., OSONO T.,  
 RA KIKUCHI H., SHIBA T., SAKAKI Y., HATORI M.,  
 RT "Genome sequence of an industrial microorganism *Streptomyces*  
 RT *avermitilis*: deducing the ability of producing secondary  
 RT metabolites.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 DR EMBL:BA000030; BAC69721.1; -; Genomic DNA.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 208 AA; 20876 MW; 4CD5B8C11AF502 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 208;  
 Best Local Similarity 16.7%; Pred. No. 2e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12  
 Db 54 CPAASSASAAAC 65

RESULT 32  
 067VV9\_ORYSA

ID 067VV9\_ORYSA PRELIMINARY; PRT; 210 AA.  
 AC 067VV9;  
 DT 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)  
 DE Hypothetical protein OSUNBA0023122.2 (Hypothetical protein  
 DE P0530H05.17).  
 GN Name=OSUNBA0023122.2; Synonym=P0530H05.17;  
 OS Oryza sativa (japonica cultivar-group),  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoideae; Oryza.  
 OC NCBI\_Taxid=39947;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.

RA Sasaki T., Matsumoto T., Yamamoto K.,  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC  
 RT clone:OSUNBA0023122.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Sasaki T., Matsumoto T., Yamamoto K.,  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC  
 RT clone:P0530H05.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL:AP004733; BAD37710.1; -; Genomic\_DNA.  
 DR EMBL:AP003541; BAD37385.1; -; Genomic\_DNA.  
 DR Gramene; 067VV9; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 210 AA; 21575 MW; C5EB3FFP3B9F4457 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 210;  
 Best Local Similarity 16.7%; Pred. No. 2e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12  
 Db 57 CSSTSSSSSRRC 68

RESULT 33  
 05K9J0\_CRYNE PRELIMINARY; PRT; 214 AA.  
 ID 05K9J0;  
 AC 05K9J0;  
 DT 10-MAY-2005 (TREMBlrel. 30, Created)  
 DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)  
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=CNK01770;  
 OS *Cryptococcus neoformans* var. *neoformans* JEC21.  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
 OC Tremellales; Tremellales; Tremellaceae; Filobasidiella.  
 OC NCBI\_Taxid=214684;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.

RC STRAIN=JEC21;  
 RA Lotfus B., Amedeo P., Roncaglia P., Vamathevan J., Uterback T.,  
 RA Van Aken S., Frazer C.,  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RC STRAIN=JEC21;  
 RX PubMed=15653466; DOI=10.1126/science.1103773;  
 RA Lotfus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,  
 RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,  
 RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,  
 RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,  
 RA Huang J.C., Jambon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,  
 RA Kwon-Chung K.J., Lengele K.B., Malt R., Marra M.A., Matra K.B.,  
 RA Mathewson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,  
 RA Schain J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,  
 RA Suh B.B., Tenney A., Uterback T.R., Wicks B.L., Wortman J.R.,  
 RA Mye N.H., Kronstad J.W., Lodge J.K., Heltman J., Davis R.W.,

RA Fraser C.M., Hyman R.W.;  
RT "The genome of the basidiomycetous yeast and human pathogen  
RT Cryptococcus neoformans.";  
RT Science 307:1321-1324 (2005).  
DR EMBL: AE017351; AAW46337.1; -; Genomic\_DNA.  
DR GO: GO:0005634; C:nucleus; IEA.  
DR GO: GO:0005732; C:small nuclear ribonucleoprotein complex; IEA.  
DR GO: GO:0006397; P:mRNA processing; IEA.  
DR InterPro: IPR001163; LSM\_snrNP\_core.  
DR Pfam: PF01423; LSM; 1.  
DR SMART; SM00651; Sm; 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 214 AA; 23828 MW; 9DFEPCED4844BCE42 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 214;  
Best Local Similarity 16.7%; Pred. No. 2e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 9 CVASSTSSVASC 20

RESULT 34  
O5SK09 CRYNE PRELIMINARY; PRT; 215 AA.  
ID O5SK09;  
AC O5SK09;

DT 13-SEP-2005 (TRENBLrel. 31, Created)  
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)  
DE Hypothetical protein.  
ORFNames=CNBK1770;

OS Cryptococcus neoformans var. neoformans B-3501A.  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.  
OK NCBI\_TaxID=283643;  
RN [1]

RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=B-3501A;  
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,  
RA Wicks B.L., Fu J., Davis R.W.;

RT "Cryptococcus neoformans serotype D sequencing.";  
RT Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.

DR EMBL; AAEY0100052; EAL18156.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 215 AA; 24014 MW; 1CEDEPCED4844BCE CRC64;

Query Match 57.1%; Score 16; DB 2; Length 215;  
Best Local Similarity 16.7%; Pred. No. 2e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 9 CVASSTSSVASC 20

RESULT 35  
O5K919 CRYNE PRELIMINARY; PRT; 215 AA.  
ID O5K919;  
AC O5K919;

DT 10-MAY-2005 (TRENBLrel. 30, Created)  
DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)  
DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)  
DE Hypothetical protein.  
ORFNames=CNK01770;

OS Cryptococcus neoformans var. neoformans JEC21.  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.  
OK NCBI\_TaxID=214684;  
RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=JEC21;  
RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Uterback T.,  
RA Van Aken S., Fraser C.; to the EMBL/GenBank/DBJ databases.  
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=JEC21;  
RX PubMed=15653466; DOI=10.1126/science.1103773;  
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,  
RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,

RA Bosdet I.B., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,  
RA D'Souza C.A., Fox D.S., Gribberg V., Fu J., Fukushima M., Haas B.J.,  
RA Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,  
RA Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,

RA Mathewson C.A., Mitchell T.G., Petrea M., Riggs F.R., Salzberg S.L.,  
RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,  
RA Suh B.B., Tenney A., Uterback T.R., Wicks B.L., Wortman J.R.,  
RA Wye N.H., Kronstad J.W., Lodge J.K., Heltman J., Davis R.W.,  
RA Fraser C.M., Hyman R.W.;

RT "The genome of the basidiomycetous yeast and human pathogen  
RT Cryptococcus neoformans.";  
RL Science 307:1321-1324 (2005).  
DR EMBL; AE017351; AAW46338.1; -; Genomic\_DNA.

DR GO: GO:0005634; C:nucleus; IEA.  
DR GO: GO:0005732; C:small nuclear ribonucleoprotein complex; IEA.  
DR GO: GO:0006397; P:mRNA processing; IEA.  
DR InterPro: IPR001163; LSM\_snrNP\_core.  
DR Pfam: PF01423; LSM; 1.  
DR SMART; SM00651; Sm; 1.  
KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 215 AA; 24014 MW; 1CEDEPCED4844BCE CRC64;

Query Match 57.1%; Score 16; DB 2; Length 215;  
Best Local Similarity 16.7%; Pred. No. 2e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 9 CVASSTSSVASC 20

RESULT 36  
O4QSJ9 GLECA PRELIMINARY; PRT; 227 AA.  
ID O4QSJ9;  
AC O4QSJ9;

DT 13-SEP-2005 (TRENBLrel. 31, Created)  
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)  
DE Putative non-reducing polyketide synthase (Fragment).  
OS Pertusaria coronata.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Lecanoromycetes;  
OC Pertusariales; Pertusariaceae; Pertusaria.  
OK NCBI\_TaxID=232756;  
RN [1]

RP NUCLEOTIDE SEQUENCE.  
RX PubMed=15927215; DOI=10.1016/j.phytochem.2005.04.014;  
RA Schmitt I., Martin M.P., Kautz S., Lumsch H.T.;  
RT "Diversity of non-reducing polyketide synthase genes in the  
RT Pertusariales (lichenized Ascomycota): A phylogenetic perspective.";

DR EMBL; AY918748; AAY00086.1; -; Genomic\_DNA.  
DR InterPro: IPR000794; Ketoacyl synth.  
DR Pfam: PF00109; ketoacyl-synth; 1.  
DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; 1.

KW Transferase.  
FT NON\_TER 227  
FT NON\_TER 227  
SQ SEQUENCE 227 AA; 24445 MW; 4D57BB346B62A8C CRC64;

Query Match 57.1%; Score 16; DB 2; Length 227;  
Best Local Similarity 16.7%; Pred. No. 2e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12  
 Db 95 CSSSFAAISTAC 106

RESULT 37  
 QJUNB5\_GIALA PRELIMINARY; PRT; 230 AA.  
 AC QJUNB5;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Variant-specific surface protein type 4 tspl1/tsa417-like  
 DE (Fragment).  
 GN Name=vsp417-4;  
 OS Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
 OC NCBI\_TaxId=5741;  
 RN (1)

RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=99229863; PubMed=10215024; DOI=10.1016/S0166-6851(98)00183-2;  
 RA E. P. L., Darby J. M., Mayrhofer G.;  
 RT "A new locus (vsp417-4) belonging to the tsa417-like subfamily of  
 RT variant-specific surface protein genes in Giardia intestinalis."  
 RL Mol. Biochem. Parasitol. 93:55-68(1999).  
 DR EMBL; U89153; AAD05041.1; -; Genomic\_DNA.  
 DR InterPro; IPR006212; Furin repeat.  
 DR InterPro; IPR005127; Giardina\_VSP.  
 DR InterPro; IPR006210; IEGF.  
 DR Pfam; PF03302; VSP; 1.  
 DR SMART; SM00181; EGF; 2.  
 DR SMART; SM00261; FU; 1.  
 FT NON\_TER 1  
 FT NON\_TER 230  
 FT NON\_TER 230  
 SQ SEQUENCE 230 AA; 23336 MW; 45E22ECD1C0084F1 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 230;  
 Best Local Similarity 16.7%; Pred. No. 2e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12  
 Db 22 CSSTVRAAATC 33

RESULT 38  
 Q8LD53\_ARATH PRELIMINARY; PRT; 234 AA.  
 AC Q8LD53;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE BHLH transcription factor, putative.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 OC Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxId=3702;  
 RN (1)

RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22088475; PubMed=12093376;  
 RA Haas B. J., Volfovsky N., Town C. D., Troukhan M., Alexandrov N.,  
 RA Feldmann K. A., Flavell R. B., White O., Salzberg S. L.;  
 RT "Full-length messenger RNA sequences greatly improve genome  
 RT annotation";  
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).  
 RN (2)

RP NUCLEOTIDE SEQUENCE.  
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (BHLH) domain.

DR EMBL; AY086197; AAM64276.1; -; mRNA.  
 DR HSSP; P36956; IAM9.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0030528; F:transcription regulator activity; IEA.  
 DR GO; GO:0045449; P:regulation of transcription; IEA.  
 DR InterPro; IPR001092; HLH\_basic.  
 DR Pfam; PF00010; HLH; 1.  
 DR SMART; SM00353; HLH; 1.  
 DR PROSITE; PS50889; HLH; 1.  
 SQ SEQUENCE 234 AA; 25470 MW; 9EBEA7D3E44F21D5 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 234;  
 Best Local Similarity 16.7%; Pred. No. 2e-07;  
 Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CXXXXXXXXXXC 12  
 Db 67 CESSATSSKAC 78

RESULT 39  
 Q9FH37\_ARATH PRELIMINARY; PRT; 234 AA.  
 AC Q9FH37;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-FEB-2003 (TrEMBLrel. 29, Last annotation update)  
 DE Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K5F14  
 DE (Hypothetical protein At5g54680; K5F14.2) (Putative bHLH transcription  
 DE factor) (Hypothetical protein At5g54680).  
 GN Name=At5g54680/K5F14.2; Synonyms=At5g54680;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 OC Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxId=3702;  
 RN (1)

RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=20181125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
 RT features of the regions of 3,076,755 bp covered by sixty p1 and TAC  
 RT clones.";  
 RL DNA Res. 7:31-63(2000).  
 RN (2)

RP NUCLEOTIDE SEQUENCE.  
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,  
 RA Palm C. J., Bower L., Jones T., Banh J., Carninci P., Chen H.,  
 RA Cheuk R., Chung M. K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
 RA Kim C., Lin J., Liu S. X., Narusaka M., Pham P. K., Sakano H.,  
 RA Sakurai T., Satou M., Seki M., Shimu P., Yamada K., Shinozaki K.,  
 RA Ecker J., Theologis A., Davis R. W.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN (3)

RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22598051; PubMed=12679534; DOI=10.1093/molbev/msg088;  
 RA Heim W. A., Jakoby M., Weber M., Martin C., Weisshar B., Bailey P. C.;  
 RT "The basic helix-loop-helix transcription factor family in plants: a  
 RT genome-wide study of protein structure and functional diversity.";  
 RL Mol. Biol. Evol. 20:735-747(2003).  
 RN (4)

RP NUCLEOTIDE SEQUENCE.  
 RA Jakoby M. J., Heim W. A., Bailey P., Martin C., Weisshar B.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RN (5)

RP NUCLEOTIDE SEQUENCE.  
 RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,  
 RA Palm C. J., Bower L., Jones T., Banh J., Carninci P., Chen H.,  
 RA Cheuk R., Chung M. K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
 RA Kim C., Lin J., Liu S. X., Narusaka M., Pham P. K., Sakano H.,  
 RA Sakurai T., Satou M., Seki M., Shimu P., Yamada K., Shinozaki K.,  
 RA Ecker J., Theologis A., Davis R. W.;



RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.  
DR EMBL; AB022214; BAB09934.1; -; Genomic\_DNA.  
DR EMBL; AY054585; AAK96776.1; -; mRNA.  
DR EMBL; AF488629; AAM10964.1; -; mRNA.  
DR EMBL; BT002189; AAM72200.1; -; mRNA.  
DR HSSP; P36956; 1AM9.  
DR GO; 0045449; P:regulation of transcription; TAS.  
DR InterPro; IPR001092; HLH\_basic.  
DR Pfam; PF00010; HLH; 1.  
DR SMART; SM00353; HLH; 1.  
DR PROSITE; PS50888; HLH; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 234 AA; 25488 MW; 987EA114244F21D1 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 234;  
Best Local Similarity 16.7%; Pred. No. 2e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
DB 67 CESSATSSKAC 78

RESULT 40  
ID 016421 CAEEL PRELIMINARY; PRT; 244 AA.  
AC 016421;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Hypothetical protein T05B4.12.  
GN ORFNames=T05B4.12;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=Bristol N2;  
RG MEDLINE=99069613; PubMed=9851916;  
RG The C. elegans sequencing consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL; AF016445; AAC69052.1; -; Genomic\_DNA.  
DR PIR; T31838; T31838.  
DR Ensemble; T05B4.12; Caenorhabditis elegans.  
DR Wormbase; WBGene00020243; T05B4.12.  
DR WormPep; T05B4.12; CE13215.  
DR InterPro; IPR003582; ShKT.  
DR Pfam; PF01549; ShTK; 3.  
DR SMART; SM00254; ShKT; 4.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 244 AA; 25347 MW; 36362492AC6B1C9 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 244;  
Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
DB 203 CTSYADSSSTSC 214

RESULT 41  
ID 0966B6 CAEEL PRELIMINARY; PRT; 247 AA.  
AC 0966B6;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Hypothetical protein.

GN ORFNames=Y46H3D.8;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=Bristol N2;  
RG MEDLINE=99069613; PubMed=9851916;  
RG The C. elegans sequencing consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL; AC006777; AAK7311.1; -; Genomic\_DNA.  
DR Ensemble; Y46H3D.8; Caenorhabditis elegans.  
DR InterPro; IPR003582; ShKT.  
DR Pfam; PF01549; ShTK; 3.  
DR SMART; SM00254; ShKT; 4.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 247 AA; 25215 MW; FF3CD3BF00DDE5A7 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 247;  
Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
DB 206 CTSYADSSSAC 217

RESULT 42  
ID 016424 CAEEL PRELIMINARY; PRT; 248 AA.  
AC 016424;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Hypothetical protein T05B4.3.  
GN ORFNames=T05B4.3;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=Bristol N2;  
RG MEDLINE=99069613; PubMed=9851916;  
RG The C. elegans sequencing consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL; AF016445; AAC69054.1; -; Genomic\_DNA.  
DR PIR; T31841; T31841.  
DR Ensemble; T05B4.3; Caenorhabditis elegans.  
DR Wormbase; WBGene00020237; T05B4.3.  
DR WormPep; T05B4.3; CE13197.  
DR InterPro; IPR003582; ShKT.  
DR Pfam; PF01549; ShTK; 3.  
DR SMART; SM00254; ShKT; 4.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 248 AA; 25659 MW; 79B281AFBBD3AFC CRC64;

Query Match 57.1%; Score 16; DB 2; Length 248;  
Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
DB 207 CTSFADSSSSC 218

RESULT 43  
ID 09XV90 CAEEL

ID Q9XV90\_CABEL PRELIMINARY; PRT; 257 AA.  
AC Q9XV90;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical protein Fl6H6.3;  
GN Name=Fl6H6.3; ORFName=Fl6H6.3;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
OC Rhabditidae; Pelodexinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=Briscot N2;  
RL Substhe L.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=Briscot N2;  
RX MEDLINE=99069613; PubMed=9851916; DOI=10.1126/science.282.5396.2012;  
RG The C. elegans sequencing consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology."  
RT Science 282:2012-2018(1998).  
RL EMBL: Z81506; CAB04129.1; -; Genomic\_DNA.  
DR PIR: T21029; T21029.  
DR Ensembl; Fl6H6.3; Caenorhabditis elegans.  
DR Wormbase; WBGene0000893; Fl6H6.3.  
DR WormPep; Fl6H6.3; CE18599.  
DR InterPro; IPR003582; ShKT.  
DR Pfam; PF01549; ShKT; 3.  
DR SMART; SM00254; ShKT; 4.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 257 AA; 27646 MW; CA204F1A6C00EB87 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 257;  
Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12  
DB 195 CTNATTTTQASC 206

RESULT 44  
08CAZ5\_MOUSE PRELIMINARY; PRT; 262 AA.  
ID 08CAZ5\_MOUSE  
AC 08CAZ5;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Mus musculus 6 days neonate skin cDNA, RIKEN full-length enriched library, clone:A030004L08 product:hypothetical protein, full insert sequence. (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Eumarchontiformes; Glires; Rodentia; Sciurognathi;  
OC Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Skin;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
CArlanci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Skin;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Caavaant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuenl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Bash G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Holtman M., Hume D.A., Kamlya M., Lee N.H.,  
RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Williams L.,  
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Skin;  
RA The FANTOM Consortium,  
RA the RIKEN genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Skin;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/9r.145100;  
CArlanci P., Shibata Y., Hayatsu N., Sugihara Y., Shibata K., Itoh M.,  
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Skin;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/9r.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Saeki N., Carninci P.,  
RA Kono H., Akiyama J., Nishi K., Katsunai T., Taahiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa K., Izawa M., Ohara E., Wachihi M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-Format sequencing pipeline with 384 multiplexed sequencer."  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Skin;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hasegawa T., Harada A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hirao T., Hirozane T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ono M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takehashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Taya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC037168; BAC29730.1; -; mRNA.  
DR Ensembl; ENSMUSG0000050239; Mus musculus.  
KW Hypothetical protein.  
FT NON TER 1  
SQ SEQUENCE 262 AA; 29111 MW; 8D441316CD4432E8 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 262;  
Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
DB 139 CTTATRSASKAC 150

## RESULT 45

ID Q7PRO7 ANOGA PRELIMINARY; PRT; 269 AA.  
AC Q7PRO7;  
DT 01-MAR-2004 (TEMBLrel. 26, Created)  
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)  
DE ENSANGP0000001657 (Fragment).  
GN ORFNames=ENSANG0000001387;  
OS Anopheles gambiae str. PEST.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;  
OX Anophelinae; Anopheles.  
[1] NCBI\_TaxID=180454;  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=PEST;  
RG The Anopheles gambiae Sequence Committee;  
RT "Anopheles gambiae re-annotation."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
[2]  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=PEST;  
RG The Anopheles gambiae Sequence Committee;  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AAA01008847; FA06779.3; -; Genomic\_DNA.  
FT NON\_TER 1 1  
FT NON\_TER 269 269  
SQ SEQUENCE 269 AA; 24228 MW; 25BBFF71FD71F1F2 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 269;  
Best Local Similarity 16.7%; Pred. No. 2.1e-07;

Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
DB 114 CASSCSTSSC 125

## RESULT 46

ID Q6ZUB7 HUMAN PRELIMINARY; PRT; 278 AA.  
AC Q6ZUB7;  
DT 05-JUL-2004 (TEMBLrel. 27, Created)  
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)  
DE Hypothetical protein FLJ43840.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;  
OC Homo  
OX NCBI\_TaxID=9606;  
RN NUCLEOTIDE SEQUENCE.  
RP TISSUE=Testis;  
RC Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,  
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,  
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
RA Kawai-Hio Y., Saito K., Sekine K., Kinouchi H., Kanda K., Wagatsuma M.,  
RA Matsumoto K., Nakamura Y., Nishikawa T., Kimura K., Yamashita H.,  
RA Murakami K., Sugihara K., Takahashi-Fujii A., Ohtsuka A., Sugiyama A.,  
RA Kawakami B., Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K.,  
RA Isogai T.;  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK125828; BAC6309.1; -; mRNA.

SQ SEQUENCE 278 AA; 27768 MW; B0C6757FC1DAE2C0 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 278;  
Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
DB 35 CTTAGSSSLSTC 46

## RESULT 47

ID Q8PPE6 XANAC PRELIMINARY; PRT; 285 AA.  
AC Q8PPE6;  
DT 01-OCT-2002 (TEMBLrel. 22, Created)  
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)  
DE Endonuclease.  
GN OrderdiocName=XAC4037;  
OS Xanthomonas axonopodis (pv. citri).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xanthomonas.  
OX NCBI\_TaxID=92829;  
RN NUCLEOTIDE SEQUENCE.  
RP STRAIN=306 / ATCC 13902 / XV 101;  
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/117459a;  
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R.,  
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,  
RA Almeida N.F., Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,  
RA Camargo L.E.A., Gamariote G., Camavan F., Cardoso J., Chambergo F.,  
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,  
RA El-Dorri H., Faria U.B., Ferreira A.U.B., Ferreira R.C.C., Gruber A.,  
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
RA Martins E.C., Meidants J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
RA Pereira H.A., Rossi A., Sene J.A.D., Silva C., de Souza R.F.,  
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
RA Setubal J.C., Kitajima J.P.;  
RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
RT host specificities."  
RU Nature 417:459-463(2002).  
DR EMBL; AF012052; AAM3872.1; -; Genomic\_DNA.  
DR HSSP; P13717; 1G8T.  
DR GO; GO:0004519; P:endonuclease activity; IEA.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro: IPR001604; Endonuclease.  
DR Pfam: PF01223; Endonuclease\_NS; 1.  
DR SMART; SM00477; NUC; 1.  
DR PROSITE; PS01070; NUCLEASE\_NON\_SPEC; UNKNOWN\_1.  
KW Complete proteome; Endonuclease.

SQ SEQUENCE 285 AA; 30358 MW; 5581A0B49A2057BB CRC64;

Query Match 57.1%; Score 16; DB 2; Length 285;  
Best Local Similarity 16.7%; Pred. No. 2.1e-07;  
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 CXXXXXXXXXXC 12  
DB 51 CTTVAATATRC 62

## RESULT 48

ID Q6ZCO1 HUMAN PRELIMINARY; PRT; 286 AA.  
AC Q6ZCO1;  
DT 05-JUL-2004 (TEMBLrel. 27, Created)  
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)

```

DE Hypochemical protein FLJ46328.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RA Niinomiya K., Wagatsuna M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa K., Yamazaki M., Sugiyama T., Irie R.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki A., Ishii S., Yamamoto J., Isono Y.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hiro Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanohori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugeno S., Nagahari K., Maunho Y., Nagai K., Isogai T.,
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128833; BAC87633.1; -; mRNA.
SQ SEQUENCE 286 AA; 30415 MW; B6CEAPFA76737B8C CRC64;

Query Match 57.1%; Score 16; DB 2; Length 286;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 231 CTAAGSSSLSTC 242

RESULT 49
Q9F934 BIFAD PRELIMINARY; PRT; 287 AA.
AC Q9F934;
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Sugar binding protein Sbp.
GN Name=sbp;
OS Bifidobacterium adolescentis.
OC Bacteria; Actinobacteria; Actinobacteriales; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=1680;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=INT57;
RA Park M.S., Yoon H.J., Ji G.E.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF213175; AAC31696.1; -; Genomic_DNA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR006059; SBP_bac_1.
DR Pfam; PF01547; SBP_bac_1; 1.
SQ SEQUENCE 287 AA; 32054 MW; 11B9382AE39B86DB CRC64;

Query Match 57.1%; Score 16; DB 2; Length 287;
Best Local Similarity 16.7%; Pred. No. 2.1e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 48 CAAIMASAAAC 59

RESULT 50
Q89DY5 BRAJA PRELIMINARY; PRT; 297 AA.
AC Q89DY5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE B117302 protein.
GN OrderedLocNames=b117302;

```

```

OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=2248498; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamideawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Ideawa K., Iritsugu M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BACS2567.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 297 AA; 32798 MW; E8B4D21866645601 CRC64;

Query Match 57.1%; Score 16; DB 2; Length 297;
Best Local Similarity 16.7%; Pred. No. 2.2e-07;
Matches 2; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CXXXXXXXXXXC 12
DB 250 CSATHSDDAAC 261

```

Search completed: January 4, 2006, 16:09:56  
 Job time : 79.3652 secs